Dissecting Melanoma Heterogeneity by Integrative Genomic Analysis for Tailored Anti-Cancer Therapy

Thesis

Dissecting Melanoma Heterogeneity by Integrative Genomic Analysis for Tailored Anti-Cancer Therapy

Matteo Dugo

OU personal identifier: C7256656

Thesis presented for the Degree of Doctor of Philosophy

The Open University, Milton Keynes (UK)

Faculty of Science Technology Engineering and Mathematics (STEM)

Discipline: School of Life, Health and Chemical Sciences

Date of submission: September 2018

Affiliated Research Centre:

Fondazione IRCCS Istituto Nazionale dei Tumori, Milan (Italy)

Director of studies: Dr Silvana Canevari

Internal supervisor: Dr Marialuisa Sensi

External supervisor: Dr Florian Markowetz

TABLE OF CONTENT

ABSTRACT

Cutaneous melanoma is a highly aggressive disease resistant to conventional treatment and characterized by poor prognosis. Targeted therapies against MAPK pathway and immune checkpoint inhibitors have dramatically improved survival of metastatic melanoma patients but the extent and duration of response are variable. Classification based on gene expression profiling have so far allowed identification of melanoma subtypes with distinctive biological features and with potential clinical impact. However, clinical translation of molecular subtypes is hampered by inconsistencies among the different classifications.

Here, through a harmonized bioinformatic analysis of public transcriptomic datasets, we compared and combined nine published classification systems to derive the consensus transcriptional subtypes of melanoma. Beyond confirming previously reported subtypes, our approach enabled the identification of a novel highly mitotic, chromosomally unstable group of melanomas that recapitulated a transitory state from a proliferative, melanocytic, differentiated phenotype to a more mesenchimal invasive program. We provided evidence that this classification has a prognostic role in metastatic melanoma patients, independently from the levels of tumour immune infiltration. We translated consensus subtypes to *in vitro* melanoma cell lines and combining them with pharmacological data we highlighted subtype-specific sensitivity to MAPK inhibitors and other drugs. Analysis of baseline gene expression data of metastatic melanoma patients treated with MAPK or immune checkpoint inhibitors showed that the predictive role of consensus subtypes in clinical setting remains to be elucidated. Finally, through the analysis of multi-omics data from the same set of patients, we comprehensively characterized the consensus subtypes at the genomic, transcriptional, and epigenomic levels.

Our results showed that melanoma gene expression classifications converged on five biological entities determined by transcriptional and epigenetic events, and with potential implications for prognostication.

DECLARATION OF AUTHORSHIP

I hereby certify that the thesis I am submitting is my own original work. All the data used in this thesis are publicly available and were properly acknowledged as references.

LIST OF ABBREVIATIONS

AJCC: American Joint Committee on Cancer

ANOVA: analysis of variance

ASR: age-standardized rate

AUC: area under curve

CDF: cumulative distribution function

CCLE: Cancer Cell Line Encyclopedia

CIN: chromosomal instability

CLND: completion lymph node dissection

cpm: counts per million

CR: complete response

CSD: chronically sun-damaged

CTdb: cancer-testis database

CTRP: Cancer Therapeutics Response Portal

EBI: European Bioinformatics Institute

ESC: embryonic stem cells

FC: fold change

FDA: Food and Drug Administration

FDR: false discovery rate

fgsea: fast gene set enrichment analysis

FPKM: fragment per kilobase million

frma: frozen robust multi-array average

GEO: gene expression omnibus

GO: gene ontology

GSEA: gene set enrichment analysis

GSVA: gene set variation analysis

GTF: general transfer format

HDAC: histone deacetylase

IPA: Ingenuity Pathway Analysis

KEGG: Kyoto Encyclopedia of Genes and Genomes

MAPK: mitogen-activated protein kinase

MCSM: melanoma clinical samples from microarrays

miRNA: microRNA

mm: millimeters

MSigDB: molecular signature database

NCBI: National Center for Biotechnology Information

NTP: nearest-template prediction

PAM: partition around medoids

PD: progressive disease

PR: partial response

RECIST: Response Evaluation Criteria In Solid Tumors

rma: robust multi-array average

ROS: reactive oxygen species

RPPA: reverse phase protein array

RSEM: RNA-Seq by expectation maximization

RSN: robust spline normalization

RTK: receptor tyrosin kinase

SKCM: skin cutaneous melanoma

SLN: sentinel lymph node

TCGA: The Cancer Genome Atlas

TMM: trimmed mean of M-values

UV: ultra-violet

WT: wild type

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

1.1 Melanoma epidemiology and risk factors

Melanoma is a relatively uncommon cancer but its worldwide incidence has been increasing faster than any other cancer type in the last 50 years (Ali et al., 2013). For 2018, approximately 287,723 new melanoma cases and 60,712 melanoma-related deaths are estimated (Bray et al., 2018). Although melanoma accounts for only 5% of all dermatological malignancies, its highly metastatic potential makes it the most lethal form of skin cancer. Incidence rates vary widely across countries and these variations are related to differences of complexion across ethnic groups and patterns of sun exposure. Melanoma is more frequent in the Caucasian population with light skin complexion due to lower melanin production and consequently decreased photoprotection. On the other hand, darker-pigmented individuals have a higher tanning ability and higher protection against ultraviolet- (UV-)A and UV-B radiation (Brenner and Hearing, 2008). Within the same ethnic group, melanoma incidence varies by geographic location (Figure 1.1) as many factors, like altitude, latitude and cloud cover, can influence UV radiation intensity. The highest incidence worldwide is observed in New Zealand and Australia with an annual age-standardized rate (ASR) of 33.6 and 33.3 cases per 100,000 inhabitants (http://gco.iarc.fr/today/). In Europe, Norway (29.6:100,000 cases per year), Denmark (27.6:100,000 cases per year), and Sweden (24.7:100,000 cases per year) have the highest ASR (http://gco.iarc.fr/today/). Southern countries, like Italy (12.4:100,000 cases per year) and Spain (6.4: 100,000 cases per year), have a lower ASR compared to Scandinavian countries. This difference is partially related to increased protection against UV radiation in darker skinned people from Southern Europe compared to fair skin people typical of Scandinavian countries.

In relation to age, melanoma affects young and middle-aged people with a median age at diagnosis of 57 years. Worldwide, melanoma ASR increases linearly with age, even if in the United States the highest incidence rate is observed in the sixth decade of life and then decreases (Matthews et al., 2017). Melanoma incidence is also related to sex, with men being in general more affected than women (Matthews et al., 2017). When sex and age are considered together it is observed that women develop melanomas at younger age, while men are more affected after the age of 40 (Rastrelli et al., 2014). This may be

partially explained by a more frequent usage among young women of artificial tanning that increases the risk of melanoma (Zhang et al., 2012).

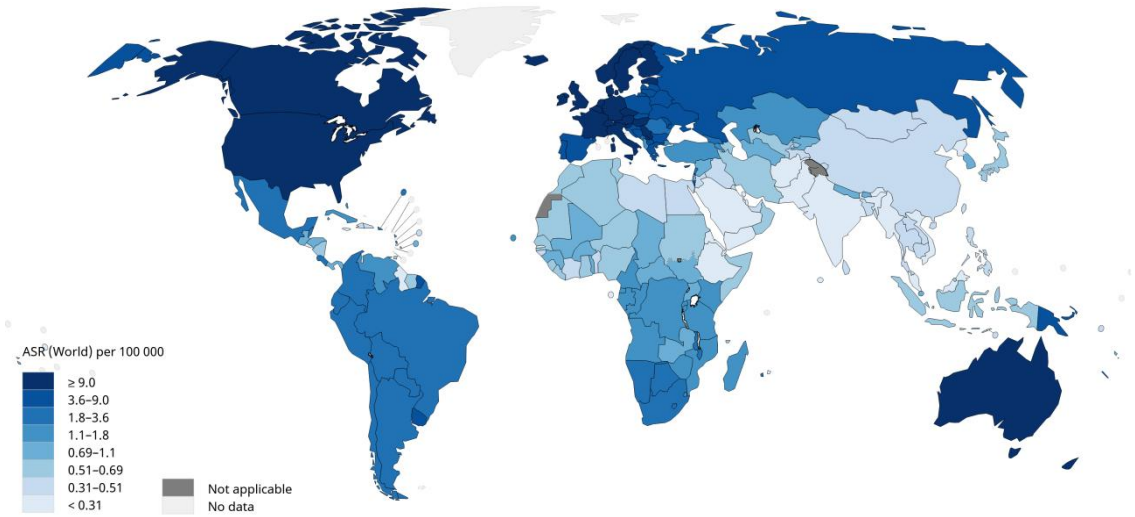

Figure 1.1. Geographical distribution of estimated age-standardized incidence rates for skin melanoma. Both sexes and all ages are considered. Source: http://gco.iarc.fr/today/.

Epidemiological studies have identified several environmental and individual melanoma risk factors, and their interplay explains the high variation in disease demographics. The most important environmental risk factor is UV radiation and especially UV-B levels. The timing and pattern of sun exposure is relevant in determining the risk of developing melanoma during the life span. A number of epidemiological studies have shown that intense and intermittent UV exposure is associated to a higher risk of melanoma (Gilchrest et al., 1999; Elwood and Jopson, 1997). In addition, intermittent sun exposure and sunburns in childhood and adolescence increase the risk of melanoma development (Whiteman et al., 2001; Holman and Armstrong, 1984). Exposure to artificial UV-A radiation has been linked to the development of melanoma. Individuals using sunbeds, especially in the young age, or receiving psoralen-UV-A radiation photochemotherapy for the treatment of psoriasis have an increased risk of melanoma (Stern, 2001; Cust et al., 2011; Veierod et al., 2010). No other environmental risk factors, including tobacco consumption (Merimsky and Inbar, 1998), have been associated with melanoma risk (Ali et al., 2013).

Melanoma susceptibility is increased by individual risk factors including congenital or acquired melanocytic nevi, familial history of melanoma, genetic susceptibility and previous melanoma diagnosis. The risk of melanoma is positively correlated with the number and size of melanocytic nevi (Holly et al., 1987; Grob et al., 1990). About 25% of melanomas, indeed, arise from a pre-existing nevus (Bevona et al., 2003). A family history

of melanoma constitutes a strong risk factor for the disease and is linked to the transmission of predisposing genetic mutations (Goldstein and Tucker, 2001). Germline mutations in the two suppressor genes CDKN2A and CDK4 are the two most frequent alterations found in individuals from melanoma-prone families (Gruis et al., 1995; Tsao and Niendorf, 2004). Additional inherited diseases including xeroderma pigmentosum, Li-Fraumeni cancer syndrome, familial retinoblastoma and Lynch syndrome type II confer higher risk of melanoma (Goldstein and Tucker, 2001).

Individuals with reduced protection from UV radiation due to fair skin, red hair and light eyes have a higher risk of developing melanoma (Titus-Ernstoff et al., 2005). This phenotype is conferred by the presence of loss of function polymorphisms in the melanocortin 1 receptor (MC1R) gene. These polymorphisms impair the synthesis of melanin and confer reduced ability to tan and photo-protection (Sturm, 2002; Kanetsky et al., 2006). A personal history of melanoma increases the chance of developing further primary melanomas (Levi et al., 2005; DiFronzo et al., 1999).

1.2 Melanoma staging and prognostic stratification

Staging is based on the TNM classification system where T refers to the features of the primary tumour, N to the number of nearby lymph nodes affected by cancer and M to the presence of metastases. Guidelines for melanoma staging are delineated by the American Joint Committee on Cancer (AJCC) with the last updated version published in 2017 (eighth edition) (Gershenwald et al., 2017).

The TNM classification provides a detailed description of the disease. The T-category stratification is defined by two independent prognostic factors: tumour thickness and ulceration (**Table 1.1**). Tumour thickness, also known as Breslow index (Breslow, 1970), is associated with worse outcome (Balch et al., 2009). Ulceration is a condition in which the dermis overlying the primary tumour is not intact according to microscopic evaluation of the tissue. The presence of ulceration is associated to a more aggressive phenotype and worse prognosis (Balch et al., 1980; McGovern et al., 1982). The number of lymph nodes affected by metastases defines the N-category that is further stratified in “a”, “b”, or “c” depending on the regional tumour burden (presence of clinically occult or clinically detected metastases) or presence of in-transit or satellite metastases (**Table 1.1**). The site of the metastases is used to stratify patients in four M1 subgroups, from “a” to “d” (**Table 1.1**).

TNM classification is used for prognostic stratification of patients in five stages with corresponding subgroups (**Table 1.2**). Stage 0 melanomas are known as melanoma *in situ* and are the earliest form of melanoma. Tumour cells are confined to the epidermis and have not spread to the surrounding tissue, lymph nodes or distant sites. Patients with melanoma *in situ* have a low risk of developing a local recurrence or metastases and surgery is curative in the majority of cases. Stage I and stage II melanomas are localized tumours that have grown deeper in the skin but have not spread to the lymph nodes or to distant sites. Patients with localized melanoma are usually cured by surgery and have a 5-year survival rate of 98.4% (Noone et al., 2017). Stage III melanomas are tumours that have spread to regional lymph nodes or have developed in-transit or satellite metastasis with no evidence of distant metastasis. According to the AJCC staging system, there are four stage III subgroups (A, B, C, and D) depending on primary tumour thickness and ulceration, number of lymph nodes affected by metastases, the regional tumour burden, and the presence of in-transit or satellite metastases (Gershenwald et al., 2017) (**Table 1.1**). Similarly to localized melanomas, tumour thickness and ulceration of the primary tumour are independent predictors of worse survival for patients with stage III melanoma (Gershenwald et al., 2017). The survival rate of patients with stage III disease is heterogeneous with five-year melanoma-specific survival ranging from 93% for patients with stage IIIA melanoma to 32% for stage IIID disease (Gershenwald et al., 2017). Stage IV melanomas are associated with metastases beyond regional lymph nodes that have reached distant sites (M1). The 5-year survival of stage IV melanoma patients is 22.5% (Noone et al., 2017). Another important prognostic factor for stage IV melanoma is the serum level of lactate dehydrogenase (LDH) that adds a further designation of the M-category. According to LDH, patients are assigned to the 0 group (not elevated LDH) or 1 group (elevated LDH) with the latter defining a decreased survival rate among stage IV patients.

In addition to the AJCC staging system other factors such as age at diagnosis, sex, primary tumour site and number of distant metastases have been shown to be of prognostic relevance. Older age at diagnosis is generally associated to worse primary tumour features, such as increased Breslow index and presence of ulceration, and decreased overall survival (Balch et al., 2013; Chao et al., 2004). Among melanoma patients, females have generally better outcome compared to men. The reason of this difference is not well understood but a possible explanation could be the younger age at

disease and the difference in tumour site (Joosse et al., 2012; Joosse et al., 2013). The site of primary tumour is another prognostic factor with lesions arising on the trunk and head and neck associated with worse survival despite they have a lower rate of positive sentinel lymph nodes (Fadaki et al., 2013).

Beyond traditional clinical and histological parameters molecular prognostic signatures have been identified through the use of gene expression microarrays. Numerous gene expression-based prognostic signatures have been reported in the literature (Winnepenninckx et al., 2006; John et al., 2008; Bogunovic et al., 2009; Conway et al., 2009; Jonsson et al., 2010; Mann et al., 2013), and despite their little overlap in terms of genes they were shown to converge on the role of immune system in determining a better prognosis (Schramm et al., 2012). Although gene expression signatures seem a promising tool for prognosis, their translation in the clinical use has not yet been realized due to lack of agreement between signatures, poor validation and assessment of their performance in comparison to current clinic-pathologic markers of outcome.

Table 1.1. TNM classification according to the AJCC staging system (Gershenwald et al., 2017).

T classification	
T1	
-A	<0.8 mm thickness with no ulceration
-B	0.8-1 mm or < 0.8 mm with ulceration
T2	
-A	>1-2 mm thickness with no ulceration
-B	>1-2 mm thickness with ulceration
T3	
-A	>2-4 mm thickness with no ulceration
-B	>2-4 mm thickness with ulceration
T4	
-A	>4 mm thickness with no ulceration
-B	>4 mm thickness with ulceration
N classification	
N0	No regional lymph nodes affected
N1a-c	One lymph node affected with micro-metastasis or macro-metastasis, or in-transit or satellite metastasis
N2a-c	Two or three lymph nodes affected, or at least one lymph node affected and in-transit or satellite metastasis
N3a-c	At least four lymph nodes affected , or at least two lymph node affected and matted nodes
M classification	
M0	No distant metastasis
M1	
a(0)	Distant metastasis to skin, not elevated LDH
a(1)	Distant metastasis to skin, elevated LDH
b(0)	Lung metastasis, not elevated LDH
b(1)	Lung metastasis, elevated LDH
c(0)	Visceral non-CNS metastasis, not elevated LDH
c(1)	Visceral non-CNS metastasis, elevated LDH
d(0)	CNS metastasis, not elevated LDH
d(1)	CNS metastasis, elevated LDH

Table 1.2. Pathological stage grouping according to the AJCC staging system (Gershenwald et al., 2017).

Stage	T classification	N classification	M classification
Stage 0	Tis	N0	M0
Stage 1A	T1a or T1b	N0	M0
Stage 1B	T2a	N0	M0
Stage 2A	T2b or T3a	N0	M0
Stage 2B	T3b or T4a	N0	M0
Stage 2C	T4b	N0	M0
Stage 3A	T1a, T1b, or T2a	N1a or N2a	M0
Stage 3B	T0	N1b or N1c	M0
Stage 3B	T1a, T1b, or T2a	N1b, N1c, or N2b	M0
Stage 3B	T2b or T3a	N1a-N2b	M0
Stage 3C	T0	N2b, N2c, N3b, or N3c	M0
Stage 3C	T1a-T3a	N2c-N3c	M0
Stage 3C	T3b or T4a	Any N ≥ N1	M0
Stage 3C	T4b	N1a-N2c	M0
Stage 3D	T4b	N3a/b/c	M0
Stage 4	Any T	Any N	M1

1.3 Current treatment options in melanoma

Treatment options for melanoma have changed in recent years. The choice of the treatment depends on the stage and location of the disease, the presence of specific genetic alterations affecting tumour cells, and the overall patient's health conditions.

1.3.1 Conventional treatments: surgery and chemotherapy

For patients with early-stage melanoma (stage 0-IIIB) surgical excision is the primary treatment and in most cases is curative (Ross and Gershenwald, 2011). The surgery procedure occurs through a wide local excision of the tumour with specific safety margins depending on the thickness of the melanoma. When melanoma is thicker than 1 mm or 0.8 mm with ulceration, a sentinel lymph node (SLN) biopsy can be performed to assess the dissemination of the disease to surrounding lymph nodes (Wong et al., 2018). If the SLN is positive for tumour cells, completion lymph node dissection (CLND) is recommended to remove all the remaining lymph nodes. However, different studies that evaluated the outcome of patients who did or did not undergo CLND showed no differences in survival between the two groups (Kingham et al., 2010; Leiter et al., 2016; Fioranelli et al., 2017). In stage IV melanoma surgical removal of metastases is rarely proposed if not for palliative care, but metastasectomy could bring survival benefit in carefully selected patients (Sosman et al., 2011; Wei et al., 2014; Martinez and Young, 2008).

The use of cytotoxic chemotherapeutic agents has been the earliest pharmacological treatment option for advanced melanoma. Until 2011, Dacarbazine, an alkylating agent approved in 1974 by the Food and Drug Administration (FDA) agency, was the only standard pharmacological option for patients with metastatic melanoma, but the overall response rate was < 20% with median response duration of 5-6 months, complete response < 5% of cases and a 5-year survival achieved in 2%-6% of patients (Serrone et al., 2000; Kim et al., 2010). Since FDA approval, dacarbazine has been tested in combinations with many other drugs, but no clinically meaningful improvements in survival have been observed (Eggermont and Kirkwood, 2004). Another chemotherapeutic drug used for advanced melanoma treatment is temozolomide, an oral prodrug of the active metabolite of dacarbazine. Given its small size and lipophilic properties it can cross the blood-brain barrier thus being more effective for the treatment of brain metastases (Agarwala and Kirkwood, 2000). Compared to dacarbazine,

temozolomide has similar response rate and median overall survival (Middleton et al., 2000; Patel et al., 2011).

1.3.2 Targeted therapies

Ongoing research has identified several genetic mutations affecting key signaling pathways in melanoma. This has led to the development of small molecule inhibitors that target these mutated proteins, which are important for the progression of the disease.
In 2002, it has been discovered that BRAF is the most frequently mutated gene in melanoma (Davies et al., 2002). BRAF is a serine-threonine kinase acting in the mitogen-activated protein kinase (MAPK) signaling pathway that control proliferation and survival of cells. Mutations in the BRAF gene are associated to constitutive pathway activation, resulting in enhanced growth and proliferation of cancer cells. The most common BRAF mutation is $BRAF^{V600E}$ where amino acid 600 changes from the normal valine to glutamic acid. Other less frequent mutations at this site are valine-to-lysine ($BRAF^{V600K}$) and valine-to-arginine ($BRAF^{V600R}$) substitutions.

Vemurafenib was the first BRAF-mutant inhibitor approved by FDA in 2011 for the treatment of unresectable stage III or stage IV melanoma patients harboring $BRAF^{V600E}$ mutation. Compared to dacarbazine, vemurafenib showed superior overall response rate, progression-free and overall survival in clinical trials including unresectable stage III/IV melanoma patients with $BRAF^{V600E/K}$ mutations (Chapman et al., 2011; Sosman et al., 2012). Another inhibitor of BRAF-mutant protein is dabrafenib that was approved by FDA in 2013 for the treatment of patients with unresectable or metastatic melanoma with $BRAF^{V600E/K}$ mutations. Dabrafenib showed similar beneficial effects of vemurafenib (Hauschild et al., 2012).

Inhibition of the MAPK pathway activity can be reached through the inhibition of downstream effectors of BRAF. MEK is a downstream target of BRAF and in 2013 the FDA approved the MEK inhibitor trametinib for the treatment of unresectable stage III/IV melanoma possessing $BRAF^{V600E/K}$ mutations. In BRAF-mutant metastatic melanoma, trametinib showed improved progression-free and overall survival as compared to chemotherapeutic treatment (Flaherty et al., 2012). In 2015, FDA approved another MEK inhibitor, cobimetinib, to be used in combination with the BRAF inhibitor vemurafenib for the treatment of BRAF-mutant melanomas that cannot be surgically removed or have already metastasized. This combinatorial therapy showed a significant improvement of progression-free survival compared to vemurafenib alone (Larkin et al., 2014).

Although targeted therapies have improved survival of patients with advanced melanoma, durable responses are not observed and most patients develop resistance (Trunzer et al., 2013). Furthermore, about 20% of patients lack anti-tumour response to BRAF inhibitors due to intrinsic resistance (Van Allen et al., 2014). Many mechanisms of resistance have been described. These include reactivation of the MAPK and PI3K-AKT pathways through MEK (Wagle et al., 2011; Trunzer et al., 2013) and NRAS (Nazarian et al., 2010) mutations, BRAF-mutant amplification (Shi et al., 2012), loss of NF1 (Whittaker et al., 2013) or PTEN (Paraiso et al., 2011), splicing variants of BRAF transcript (Poulikakos et al., 2011), activation of COT kinase (Johannessen et al., 2010), disrupted feedback regulation (Lito et al., 2012), receptor tyrosine kinases up-regulation (Kabbarah et al., 2010; Villanueva et al., 2010), and sustained signaling through stromal secretion of HGF (Straussman et al., 2012). In addition, intra-tumour heterogeneity can contribute to therapeutic failure. Cancer cells with different molecular alterations can coexist within the same tumour and treatment can induce a selective pressure that can enhance the expansion of subclones with features of resistance (Yancovitz et al., 2012; McGranahan and Swanton, 2017).

Another type of targeted therapy includes immune checkpoint inhibitors, antibodies that indirectly affect the tumour by targeting immune inhibitory molecules, preventing their binding to ligands and boosting the anti-tumour immune response. Melanoma is one of the most immunogenic malignancies with a high mutational burden that has been associated with response to immunotherapy in multiple cancer types. Thus, melanoma results a good candidate for this type of treatment. The first immune checkpoint inhibitor approved by FDA in 2011 was the anti-CTLA-4 monoclonal antibody ipilimumab, after reports of improved overall survival in patients with previously treated metastatic melanoma (Hodi et al., 2010). CTLA-4 is a receptor expressed on activated T-cells that, following ligand binding, down-regulates the extent of T-cell activation. (Brunet et al., 1987). Targeting PD-1, another immune checkpoint receptor, and its ligands, programmed cell death ligand 1 (PD-L1) and PD-L2, is also emerging as a promising immunotherapeutic modality (Okazaki and Honjo, 2006). Compared to targeted therapy, immune checkpoint inhibitors show higher toxicity and lower rapidity of response but they can achieve more long-lasting responses (Ratterman et al., 2016). Two anti-PD-1 monoclonal antibodies, nivolumab and pembrolizumab, were approved by FDA in 2014 for the treatment of patients with unresectable stage III or metastatic melanoma, or

patients progressing after receiving ipilimumab. Compared to ipilimumab, they provide better outcome and less toxicity (Wolchok et al., 2017; Schachter et al., 2017).

1.4 Melanocytes and melanomagenesis

Melanocytes are neural crest-derived cells present in the basal layer of epidermis and in hair follicles (Miller and Mihm, Jr., 2006). A considerable number of melanocytes can also be found in the uveal tract of the eye and at lower density in other tissues such as mucosal surfaces and meninges. The principal role of melanocytes is to provide protection from UV radiations through the production of melanin (Sturm, 1998) (**Figure 1.2**). Upon UV-induced DNA damage, skin keratinocytes secrete the α-melanocyte stimulating hormone (α-MSH) that binds MC1R on melanocytes (Swope et al., 2014). MC1R activates the transcription of micropthalmia-associated transcription factor (MITF), which in turn induces the expression of melanin synthesis genes and stimulates melanin production (Chen et al., 2014). Melanin is delivered to surrounding keratinocytes through melanosomes and protects them from further DNA damage by scattering and absorbing UV radiation.

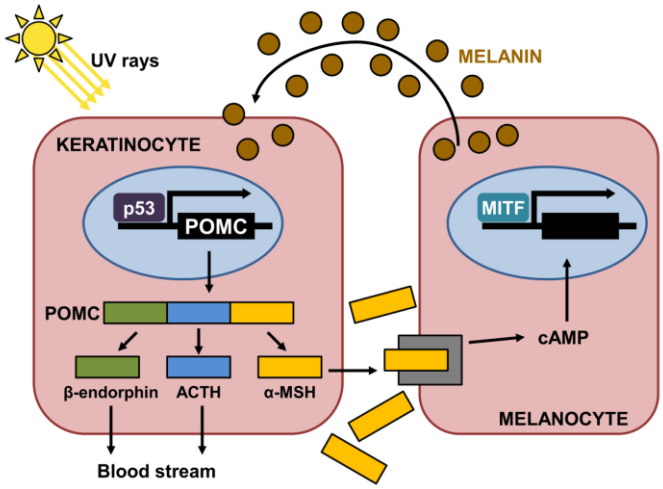

Figure 1.2. Cellular response to UV-radiation. In response to sun exposure keratinocytes synthesize the pro-opiomelanocortin (POMC) polypeptide in a p53-dependent manner. POMC is processed into its active components: α-melanocyte stimulating hormone (α-MSH), adrenocorticotropic hormone (ACTH), and β-endorphin. α-MSH binds to the MC1R receptor on melanocytes that produce melanin through the activation of pigment-related genes regulated by MITF.

Melanocytes produce two types of melanin: the black/brown eumelanin and the red pigment pheomelanin. Their ratio confers the large variety of skin pigmentation in humans. After UV absorption, pheomelanin produces reactive oxygen species (ROS) that can affect melanocytes physiological processes and can induce DNA damage if not

counterbalanced by the black eumelanin (Morgan et al., 2013). Thus, eumelanin contributes to DNA damage defense by absorption of UV radiation and ROS elimination.

The malignant transformation of melanocytes is the result of the accumulation of mutations leading to uncontrolled growth and of a deregulated interaction with the tumour microenvironment. This process follows a sequential genetic model that results in the constitutive activation of oncogenic signaling pathways. The activating $BRAF^{V600}$ mutation is the most frequent alteration found in melanomas and benign nevi, suggesting that this mutation occurs early during transformation. However, melanocytes harboring BRAF or NRAS mutations undergo oncogene-induced senescence (Michaloglou et al., 2005). This suggests that BRAF or NRAS mutations alone cannot give rise to malignant cells. Further events are required to overcome the senescent state. Progression to melanoma *in situ* requires additional mutations such as mutations in TERT promoter, mutations or loss of CDKN2A or PTEN. The TERT gene encodes telomerase reverse transcriptase, the catalytic subunit of telomerase. This enzyme plays a key role in telomere maintenance during embryonic development but it is repressed in most adult somatic cells. Re-expression of telomerase enables cancer cells to overcome the anti-proliferative effect of telomeres shortening during cell replication. CDKN2A locus encodes two proteins ($p16^{INK4A}$ and $p14^{ARF}$) that suppress cell proliferation in presence of DNA damage or oncogene activation. PTEN encodes a phosphatase that functions as a tumour suppressor by negatively regulating the AKT/PKB signaling pathway. In this phase, melanoma cells are confined to the epidermis and they still do not have metastatic properties. Surgical excision of these lesions is associated with high cure rates. Accumulation of events that disrupt the control of cell migration and adhesion leads to the vertical growth phase, where melanoma cells start to grow vertically and invade the dermis. This phase is accompanied by loss of E-cadherin and expression of N-cadherin and integrins that facilitate the remodeling of the extracellular matrix and promote invasion (Hsu et al., 1996; Sanders et al., 1999). Melanoma becomes metastatic when tumour cells dissociate from the primary tumour, invade surrounding vascular and lymphatic structures and disseminate throughout the body to form a tumour in other tissues. Usually, melanoma metastases are found first in the proximal lymph nodes while metastases at distant sites appear later. This suggests a progressive model of metastatic spreading, but evidence shows that the metastatic process can start at earlier stages from the dissemination of melanocytic nevi (Shain and Bastian, 2016; Sanborn et al., 2015).

1.5 Molecular classification of melanoma

Historically, tumour classification has been primarily based on histological features and evaluation of single biomarkers. It is generally accepted that the different clinical courses of patients with histologically identical tumours is a result of molecular differences among cancers. The advent of high-throughput technologies such as microarrays, first, and next-generation sequencing (NGS), subsequently, has revolutionized the way of looking at cancer heterogeneity making possible to obtain the molecular profile of tumours on a genome-wide scale. Gene expression profiling has been proven to be a valuable tool to identify and predict cancer subtypes (Golub et al., 1999). In melanoma, many transcriptomic studies have been reported. Bittner and colleagues (Bittner et al., 2000) published in 2000 the first study reporting a molecular classification of cutaneous melanoma. Analyzing the transcriptional profiles of metastatic melanoma biopsies and cell lines using an early cDNA microarray platform, they identified through unsupervised analysis two subtypes of melanoma with a major cluster characterized by reduced expression of genes involved in motility and invasion. This finding was further supported by an independent study that identified two subtypes of metastatic melanoma with opposite gene expression patterns (Haqq et al., 2005). With the improvement of microarray technology this classification was further investigated in numerous transcriptomic studies (**Table 1.3**). *In vitro* experiments showed that one of the two subtypes was characterized by high proliferation and low motility and capacity to invade (Hoek et al., 2006); the opposite pattern of gene expression observed in the other group translated in low proliferation and high invasiveness (Hoek et al., 2006). The proliferative subtype was characterized by Wnt/β-catenin signaling, leading to up-regulation of MITF, the master regulator of melanocytic differentiation, and its target genes such as melan-A (MLANA), tyrosinase (TYR), dopachrome tautomerase (DCT) (Hoek et al., 2006). The invasive subtype showed down-regulation of these genes and high expression of TGFβ-regulated genes that inhibit Wnt signaling, such as WNT5A, and genes involved in the remodeling of the extracellular matrix (ECM) (Hoek et al., 2006). Moreover, the invasive capacity of this subtype was associated to the up-regulation of the receptor tyrosine kinase AXL that was expressed at high levels in MITF-negative (AXL-positive) tumours and down-regulated in proliferative melanomas (AXL-negative) (Sensi et al., 2011). Through immunohistochemical staining, and more recently by single-cell RNA sequencing, it was demonstrated that the two phenotypes can coexist within the same melanoma lesions

(Eichhoff et al., 2010; Tirosh et al., 2016). Overall, these findings led to the generation of the phenotype switching model. According to this model, cells that are in a proliferative state and thus contribute to tumour growth, can respond to specific microenvironmental signals and shift toward an invasive phenotype, giving them the ability to metastasize. Once an invasive cell has migrated to another site, it can shift back to a proliferative state and give rise to a metastatic lesion (Hoek et al., 2008; Hoek and Goding, 2010; Carreira et al., 2006). The mechanism of phenotype switching not only explains the possible dynamics of the metastatic process, but it was shown to be a determinant of resistance to MEK and BRAF inhibitors with the invasive subtype being less sensitive to these drugs (Zipser et al., 2011; Konieczkowski et al., 2014; Muller et al., 2014).

Over the years, new gene signatures discriminating the two melanoma subtypes have been derived, providing new insight into the molecular mechanisms and regulatory precesses at the basis of phenotype switching (Verfaillie et al., 2015; Rambow et al., 2015). Verfaillie and colleagues (Verfaillie et al., 2015) applied unsupervised clustering to gene expression data of melanoma clinical samples and identified two main subtypes that mirrored the proliferative/invasive states observed *in vitro* (Hoek et al., 2006). They identified a more comprehensive phenotype switching-related gene signature partially overlapping with the list obtained by Widmer et al (Widmer et al., 2012). In addition, through an extensive *in vitro* and *in silico* analysis of the chromatin landscape and transcriptome of melanoma cell lines, they identified TEAD factors as the master regulators of the invasive state. Subsequently, Rambow and co-workers (Rambow et al., 2015) applied a pan-cancer bioinformatics approach to transcriptomic data of cell lines and clinical tumours to identify melanocyte lineage-specific gene and microRNA signatures of invasion and proliferation.

Additional studies showed that the dual melanoma classification can be further refined with the identification of biologically meaningful subgroup within each subtype. According to the expression of RTKs, we identified two subset of the invasive state distinguished by opposite expression of ERBB3 and EGFR (Dugo et al., 2015). In a very recent work, Tsoi and colleagues identified four subtypes of melanoma cell lines that recapitulate the different steps of the melanocytes differentiation trajectory (Tsoi et al., 2018). Two subtypes, the transitory and the melanocytic, showed features of a proliferative state such as low expression of AXL and high expression of MITF, with a stronger expression of MITF-target genes in the latter group. On the other hand, the

undifferentiated and the neural crest-like subtypes presented down-regulation of MITF and up-regulation of AXL and overlapped with the $EGFR^{high}/ERBB3^{low}$ and $EGFR^{low}/ERBB3^{high}$ invasive subtypes that we previously described (Dugo et al., 2015). In addition, they showed that differentiation of melanoma cells was negatively correlated to susceptibility to ferroptosis-inducing drugs.

Beside a better knowledge in melanoma biology, molecular subtypes can be also relevant for prognostication of melanoma patients. Gene expression analysis of 57 stage IV melanoma patients led to the identification of four subtypes associated with overall survival (Jonsson et al., 2010). These subtypes partially overlapped with the invasive/proliferative classification and showed that high tumour immune infiltration is a predictor of good outcome, as previously reported (Bogunovic et al., 2009). The Cancer Genome Atlas consortium reported a stratification of melanoma patients based on whole-exome sequencing mutational analysis (The Cancer Genome Atlas Network, 2015). According to the pattern of the most frequently mutated genes, four subtypes were defined: mutant BRAF, mutant RAS, mutant NF1 and triple wild-type (WT). Despite this genomic classification can guide the choice of the best targeted treatment, it was not associated to patient's outcome. However, in their comprehensive molecular characterization of melanoma, TCGA identified three transcriptional subtypes: keratin, MITF-low and immune. The MITF-low and keratin groups corresponded to the invasive and proliferative subtypes, respectively, with the latter being associated to the worst overall survival. The immune subtype, similarly to the high-immune group identified by Jonsson and colleagues, showed marked expression of immune-response related genes and had the best overall survival (The Cancer Genome Atlas Network, 2015). In agreement with previous studies (Hoek et al., 2006; Widmer et al., 2012; Verfaillie et al., 2015), the genomic and transcriptomic subtypes identified by TCGA were not correlated, supporting the notion that genomic alterations are not the main determinants of the transcriptional phenotype of melanoma cells.

Table 1.3. Summary of gene expression-based classification systems. The signatures are listed according to the date of publication.

Classification (Reference)	Sample source	Subtype and signature generation methods	N. of genes	Subtypes
Jonsson (Jonsson et al., 2010)	**Clinical tumours**	Unsupervised hierarchical clustering + differential expression	377	Proliferative Pigmentation High-immune Normal-like
Sensi (Sensi et al., 2011)	**Cell lines**	Genes positively or negatively correlated with AXL gene (Pearson's correlation > 0.4 or < -0.4 and p-value < 0.05)	256	AXL-negative AXL-positive
Widmer (Widmer et al., 2012)	**Cell lines**	Unsupervised hierarchical clustering + differential expression	97	Invasive Proliferative
Verfaillie (Verfaillie et al., 2015)	**Cell lines**	Unsupervised hierarchical clustering + differential expression	1412	Invasive Proliferative
Dugo (Dugo et al., 2015)	**Cell lines**	Unsupervised hierarchical clustering + "Classification to Nearest Centroids" algorithm	210	$EGFR^{high}/ERBB3^{low}$ invasive $EGFR^{low}/ERBB3^{high}$ invasive Proliferative
TCGA (Cancer Genome Atlas Network, 2015)	**Clinical tumours**	consensus hierarchical clustering with the 1,500 most variably expressed genes	1500	MITF-low Keratin Immune
Rambow (Rambow et al., 2015)	**Clinical tumours + cell lines**	Similarity core analysis	55	Invasive Proliferative
Tirosh (Tirosh et al., 2016)	**Clinical tumours**	Top-100 genes with highest Pearson's correlations with AXL or	200	AXL-program MITF-program
Tsoi (Tsoi et al., 2018)	**Cell lines**	Consensus hierarchical clustering + differential expression	275	Undifferentiated Neural crest-like Transitory Melanocytic

2 AIMS

Melanoma is a heterogeneous disease at the clinical, histopathological and molecular levels. Classification is a key topic in oncology to better stratify patients and improve the prediction of outcome and design novel therapeutic strategies. Transcriptomic studies have shown that melanoma can be classified in distinct molecular subtypes characterized by specific gene expression signatures. These classifications may integrate the current pathological classifiers to improve the understanding of melanoma biology and to assist clinicians in prognosis assessment and therapeutic decision making. However, such approaches are difficult to translate from bench to bedside due to inconsistencies among the different signatures and lack of extensive analytical and clinical validation. The merging of the different classifications into a single consensus stratification system is a crucial step to facilitate the translation of molecular subtypes into the clinical practice. At present, the published melanoma gene expression classifiers have not yet been integrated in a consensus subtyping scheme.

The research work described in this thesis has the purpose to derive a consensus molecular classification of melanoma through a meta-analysis of published gene expression-based classifiers and of publicly available transcriptomic data. Specifically, three tasks have been addressed:

- To derive and validate a new consensus molecular classification of melanoma combining nine previously published melanoma subtyping systems and to identify the biological pathways distinctive of consensus subtypes;
- To evaluate the prognostic impact of consensus subtypes and their ability to predict response to drug treatments in *in vitro* models and metastatic melanoma tumours;
- To define the molecular features characterizing the consensus subtypes through an integrative analysis of “multi-omics” data.

3 MATERIALS AND METHODS

3.1 Data retrieval

For this project we made extensive use of publicly available transcriptomic and genomic data collected from different sources. For TCGA samples, all genomic data, but somatic mutations, were downloaded from the Broad Institute Firehose web portal (https://gdac.broadinstitute.org/, data version 2016_01_28). Somatic mutation data, obtained with the MuTect2 pipeline, where downloaded from the TCGA data repository (https://portal.gdc.cancer.gov/, accession date 10/04/2018). Clinical data were downloaded from the TCGA Pan-Cancer Atlas initiative (https://gdc.cancer.gov/about-data/publications/pancanatlas, accession date 09/07/2018) and updated survival were obtained from supplementary table S1 of Liu et al (Liu et al., 2018). For meta-analysis of melanoma subtypes raw microarray data from melanoma clinical specimens and cell lines were retrieved from the NCBI Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) or EBI ArrayExpress (https://www.ebi.ac.uk/arrayexpress/) databases. Downloaded datasets are listed in **Table 3.1**. For the *in vitro* model of melanocyte differentiation (Mica et al., 2013) normalized expression data were retrieved from GEO as well, with accession number GSE45227. Fragment Per Kilobase Million (FPKM) values for treated patients profiled in the studies of Hugo et al (Hugo et al., 2015; Hugo et al., 2016) were obtained from GSE65185 and GSE78220. For treated patients included in the study of Rizos et al (Rizos et al., 2014) raw gene expression data were downloaded from GSE50509. For treated patients in the study of Kwong et al (Kwong et al., 2015), aligned bam files were downloaded from the European Genome-phenome Archive repository (https://www.ebi.ac.uk/ega/home) with accession number EGAD00001001306. For cell lines of the Cancer Cell Line Encyclopedia (Barretina et al., 2012), mutational data were downloaded from the CCLE portal (https://portals.broadinstitute.org/ccle/data). Pharmacogenomic data from the Cancer Therapeutics Response Portal (CTRP) version 2 (Seashore-Ludlow et al., 2015) were downloaded from the Cancer Target Discovery and Development data portal (https://ocg.cancer.gov/programs/ctd2/data-portal).

Table 3.1. List of public gene expression datasets of melanoma cell lines and clinical tumours used for consensus subtypes analysis. *: number of samples selected for analysis; †: obtained from the authors.

	Dataset	N. of melanoma samples*	Microarray platform
Clinical tumours (MCSM)	GSE12627	43	Affymetrix Human Gene U133 A
	GSE15605	56	Affymetrix Human Gene U133 Plus 2.0
	GSE19234	43	Affymetrix Human Gene U133 Plus 2.0
	GSE22153	57	Illumina HumanWG-6 v2
	GSE22154	22	Illumina HumanHT-12 v3
	GSE29359	76	Illumina HumanRef-8 v2
	GSE30812	41	Affymetrix Human Gene 1.0 ST Array
	GSE35640	65	Affymetrix Human Gene U133 Plus 2.0
	GSE46517	51	Affymetrix Human Gene U133 A
	GSE50493	71	Illumina HumanHT-12 v4
	GSE50509	21	Illumina HumanHT-12 v4
	GSE53118	78	Illumina HumanWG-6 v3
	GSE61992	9	Illumina HumanHT-12 v4
	GSE65904	211	Illumina HumanHT-12 v4
	GSE7553	53	Affymetrix Human Gene U133 Plus 2.0
Cell lines	E-MTAB-37	36	Affymetrix Human Gene U133 Plus 2.0
	E-MTAB-783	39	Affymetrix HT Human Gene U133 A
	GSE10916	50	Affymetrix Human Gene U133 Plus 2.0
	GSE11812	6	Affymetrix Human Gene U133 A
	GSE28335	12	Affymetrix Human Gene U133 Plus 2.0
	GSE32474	26	Affymetrix Human Gene U133 Plus 2.0
	GSE36133	57	Affymetrix Human Gene U133 Plus 2.0
	GSE4840	12	Affymetrix Human Gene U133 A+B
	GSE4843	44	Affymetrix Human Gene U133 Plus 2.0
	GSE57083	4	Affymetrix Human Gene U133 Plus 2.0
	GSE68950	43	Affymetrix HT Human Gene U133 A
	GSE7127	63	Affymetrix Human Gene U133 Plus 2.0
	GSE8332	18	Affymetrix Human Gene U133 Plus 2.0
	GSE9118	6	Affymetrix Human Gene U133 A
	Lin et al. Cancer Res 2008†	88	Affymetrix HT Human Gene U133 A
In vitro melanocyte differentiation model	GSE45227	36	Illumina HumanHT-12 v4
Anti-BRAF therapy	GSE50509	21	Illumina HumanHT-12 v4
	GSE65185	18	Illumina HiSeq 2000
	EGAD00001001306	14	Illumina HiSeq 2000
Anti-PD-1 therapy	GSE78220	26	Illumina HiSeq 2000

3.2 Gene expression data preprocessing

3.2.1 TCGA-SKCM

TCGA RNA-Seq RSEM estimated raw counts were processed using the voom pipeline from limma (Law et al., 2014). The count matrix was normalized using the Trimmed Mean of M-values (TMM) method (Robinson and Oshlack, 2010), implemented in the edgeR package (Robinson et al., 2010). TMM estimates a scaling factor used to reduce technical bias between samples due to differences in library size. Normalized data were transformed in log2 counts per million (cpm) using the limma voom function. Entrez gene IDs were converted in HUGO gene symbols using NCBI annotation (ftp://ftp.ncbi.nlm.nih.gov/gene/DATA/GENE_INFO/Mammalia/Homo_sapiens.gene_info, accession date 03/04/2018). The final dataset included 20413 unique genes and 471 samples divided in 368 metastatic melanomas and 103 primary tumours.

3.2.2 Melanoma Clinical Samples from Microarrays dataset

The Melanoma Clinical Samples from Microarrays (MCSM) dataset was assembled from 15 microarray studies obtained with Affymetrix or Illumina platforms. For each dataset raw data were normalized independently, with appropriate methods for each platform. Affymetrix gene arrays were pre-processed using the frozen robust multi-array average method (frma) (McCall et al., 2010). Affymetrix exon array were pre-processed using the robust multi-array average method (rma) (Irizarry et al., 2003). Illumina data were log2-transformed and normalized using the Robust Spline Normalization (RSN) method implemented in the lumi package (Du et al., 2008). After normalization, each dataset was annotated using Bioconductor annotation packages specific for each platform. Each normalized and annotated dataset was collapsed from probe- to gene-level using the collapseRows (Miller et al., 2011) function of the WGCNA package (Langfelder and Horvath, 2008), selecting for each gene the probe with highest variation across samples. Technical replicates and samples obtained after treatment administration were removed before the pre-processing steps. The 15 datasets were then combined at the gene-level and ComBat (Johnson et al., 2007) was applied to remove batch effects. Data quality was assessed before and after ComBat correction evaluating the pair-wise sample distance calculated as 1- Pearson's correlation coefficient. After quality control of batch-corrected data, 17 samples were identified as outliers (average correlation with all other samples < 0.7). These samples were removed from the uncorrected dataset then ComBat was newly applied. The final processed MCSM dataset consisted of 10332 unique genes and 897 samples.

3.2.3 Melanoma Cell Lines dataset

The datasets used to assemble the melanoma cell lines dataset were processed similarly to the MCSM dataset. In this case, all data were from Affymetrix Human Genome U133A-B or U133 Plus 2.0 platforms. Raw data were pre-processed using frma (McCall et al., 2010). Normalized datasets were merged at the probe-level considering only probe sets common among the different versions of the arrays. The merged dataset was collapsed to the gene-level using the collapseRows function (Miller et al., 2011) selecting for each gene the probe set with highest variation across samples. After data merging batch effects were removed using ComBat (Johnson et al., 2007). No outliers were detected after data quality control. The final dataset contained gene measurement for 12992 unique genes and 504 samples representing 327 unique melanoma cell lines.

3.2.4 Anti-MAPK and anti-PD-1 targeted therapy datasets

For GSE65185 and GSE78220 normalized FPKM values obtained by the authors of the respective publications were log2-transformed after adding a pseudo-count of 1. For GSE50509 raw Illumina microarray data were log2-transformed, normalized using the RSN method (Du et al., 2008) and collapsed at the gene-level using the collapseRows (Miller et al., 2011) function as described above. For EGAD00001001306 bam files were sorted by coordinates using samtools (Li et al., 2009) and raw gene-level counts were obtained using htseq-count (Anders et al., 2015) and Ensembl v75 GTF annotation file for hg19 assembly. Count data were filtered removing genes with less than 1 cpm in all samples, normalized using TMM (Robinson and Oshlack, 2010) and transformed in log2-cpm using voom (Law et al., 2014). For all datasets only pre-treatment samples were considered for downstream analyses.

3.3 Similarity analysis of published gene expression subtyping signatures

The original gene expression signatures of Widmer et al (Widmer et al., 2012), Sensi et al (Sensi et al., 2011), Dugo et al (Dugo et al., 2015), Verfaillie et al (Verfaillie et al., 2015), Rambow et al (Rambow et al., 2015), Tirosh et al (Tirosh et al., 2016), Jonsson et al (Jonsson et al., 2010) and Tsoi et al (Tsoi et al., 2018) were retrieved from the corresponding publications. The TCGA signature was downloaded from Broad Firehose data portal and each gene was assigned to the class with highest average expression according to the subtype assignments in the supplementary table 1D of the TCGA publication (The Cancer Genome Atlas Network, 2015). The pair-wise similarity of the gene content of the signatures was calculated using the Jaccard index that is a statistic used to evaluate the similarity between two categorical lists and is calculated as the ratio of the intersection and the union of the two lists. It ranges from 0 (no overlap) to 1 (perfect overlap). The similarity at the pathway level was calculated in the same manner, using the lists of significantly enriched canonical pathways (Benjamini-Hochberg false discovery rate (Benjamini and Hochberg, 1995) (FDR) < 0.05) identified through the use of Ingenuity Pathway Analysis (IPA, QIAGEN Inc., https://www.qiagenbio-informatics.com/products/ingenuity-pathway-analysis).

3.4 Identification of consensus melanoma subtypes

Samples included in TCGA-SKCM, MCSM and MCL datasets were classified according to the nine gene expression-based signatures using the Nearest Template Prediction (NTP) algorithm (Hoshida, 2010) implemented in the GenePattern suite (Reich et al., 2006). For each classifier samples with an NTP FDR < 0.25 were assigned to the corresponding subtype while the others were labeled as undetermined. The labels obtained from the nine classifiers were merged in a "samples x labels" matrix that was used to calculate the similarity between samples according to the Hamming distance. This kind of distance is the number of different subtype labels between two samples. If two samples are assigned to the same subtypes by the nine classifiers the Hamming distance is 0. If two samples have completely different subtypes assignments the Hamming distance is equal to the number of classifiers used, nine in this case. To identify a consensus of the multiple classifications we clustered the Hamming similarity matrix using consensus clustering (Monti et al., 2003) with the partition around medoids (PAM) algorithm, implemented in the ConsensusClusterPlus package (Wilkerson and Hayes, 2010). Consensus clustering is a resampling technique that runs the same clustering algorithm for a specified number of iterations on a random subset of the data and then provides method to identify a consensus clustering across the multiple runs. The procedure was repeated for varying number of clusters K (K = 2 to K = 10). The optimal number of clusters was identified by examination of consensus matrices, the cumulative distribution function (CDF) plot and the delta area under the CDF (ΔCDF) plot. The consensus matrix reports the proportion of times in which two samples are clustered together across the resampling runs. The CDF plot shows the CDF of the consensus matrix values at each K. The K at which the CDF reaches a maximum is the number of clusters that gives the best separation of the data. The ΔCDF plot shows the relative change in area under the CDF curve between K and K-1 to determine at which K there is no appreciable increase in the consensus.

The overlap between consensus clusters and labels from the nine classification systems was measured by Jaccard index and the statistical significance of the overlap was assessed by hypergeometric test. P-values were corrected for multiple testing using the Benjamini-Hochberg FDR (Benjamini and Hochberg, 1995). Only overlaps with Jaccard index ≥ 0.25 and FDR < 0.05 were considered significant and they were visualized through networks using Cytoscape v3.6.1.

3.5 Survival analysis

Survival analysis was performed considering only stage III/IV metastatic melanoma patients from TCGA-SKCM dataset. Cox regression analysis, Kaplan-Meier curves and log-rank test were performed using the survival R package. For univariate and multivariate Cox regression model the variables considered were consensus subtype and immune status according to CIBERSORT p-value.

3.6 Analysis of pharmacogenomic data

Consensus subtype labels for the 57 cell lines included in CCLE were extrapolated from the melanoma cell lines dataset. Association between consensus subtypes and drug sensitivity area under the fitted dose response curve (AUC) values was assessed using ANOVA followed by Tukey's post hoc correction. A $p < 0.05$ after Tukey's post hoc correction was set as significance threshold. Analysis was restricted to $BRAF^{V600E}$ mutant cell lines for selective BRAF inhibitors only.

3.7 Analysis of treated patient's datasets

Gene expression data were analyzed using the moderated t-test implemented in the limma package to identify genes whose expression was positively or negatively associated to therapeutic response. For datasets including patients receiving BRAF targeted therapy (GSE50509, GSE65185 and EGAD00001001306) expression data were analyzed in association to best overall response, defined according to Response Evaluation Criteria In Solid Tumors (RECIST) criteria (Eisenhauer et al., 2009) as reported in the relative publications (Rizos et al., 2014; Hugo et al., 2015; Kwong et al., 2015). For the anti-PD-1 dataset (GSE78220) we compared the three response categories (complete response, partial response and progressive disease) in a pair-wise fashion. The classical limma pipeline was applied for Illumina microarray data of GSE50509. For EGAD00001001306 limma was applied after TMM normalization and voom transformation. For GSE65185 and GSE78220 we applied the classical limma pipeline to FPKM values, with the trend parameter of the eBayes function set to true. For each dataset Gene Set Enrichment Analysis (GSEA) (Subramanian et al., 2005) was performed on the list of genes ranked according to the t-statistics of limma. For GSEA, consensus subtypes gene sets were obtained from TCGA-SKCM dataset, selecting genes exclusively up-regulated in each subtype at an FDR < 0.05. Single-cell melanoma specific gene sets

were obtained from supplementary table 2D of Rambow et al (Rambow et al., 2018). HALLMARK gene sets were downloaded from MSigDB database (Liberzon et al., 2015).

3.8 Molecular correlative analyses with consensus subtypes

The molecular differences characterizing each consensus subtypes were investigated in TCGA-SKCM samples exploiting all multi-omics data available. Each platform was analyzed using specific methods detailed in the next sections. The comparisons were made comparing each subtype with the others in a one-versus-all fashion, unless otherwise specified. For all analyses p-values were adjusted for multiple comparisons using the Benjamini-Hochberg FDR. An FDR < 0.05 was considered statistically significant. Association between continuous numerical variables and consensus subtypes was statistically assessed by means of Analysis of Variance (ANOVA), followed by Tukey's post-hoc test. Association between categorical variables and consensus subtypes was performed using Fisher's exact test.

3.8.1 Transcriptomics

Differentially expressed genes were identified using the voom limma pipeline (Law et al., 2014). Pathways associated to each subtype were identified using the Gene Set Variation Analysis (GSVA) method implemented in the GSVA package (Hanzelmann et al., 2013). This method estimates the enrichment of pathways for each single sample and enables the comparison of pathway enrichment scores across the samples of an expression dataset. For GSVA we used gene sets of the C5BP Gene Ontology (GO) biological process collection downloaded from the MSigDb database v6.1 (Liberzon et al., 2011). The top-10% associated pathways in each subtype were then represented using the EnrichmentMap app (Merico et al., 2010) and Cytoscape v3.6.1. The EnrichmentMap algorithm generates a network connecting gene sets with genes in common. A threshold of 0.6 for the combined Jaccard and Overlap coefficients was applied to connect gene sets. The obtained network was manually curated to group clusters of gene sets in functional categories. Unconnected nodes were removed from the network. Additional custom gene sets were tested: Widmer and Verfaillie invasive/proliferative signatures (Widmer et al., 2012; Verfaillie et al., 2015), Sensi AXL+/- signatures (Sensi et al., 2011), Tirosh MITF/AXL signatures (Tirosh et al., 2016) and Tirosh cell type-specific genes (Tirosh et al., 2016).

GSEA was performed on the list of genes pre-ranked according to the t-statistics obtained with limma. GSEA software was obtained from the Broad Institute website (http://software.broadinstitute.org/gsea/index.jsp). The cancer-testis antigens gene set was obtained from the Cancer-Testis database (CTdb, http://www.cta.lncc.br/) (Almeida et al., 2009). The chromosomal instability signature was obtained from Carter et al (Carter et al., 2006).

Association between consensus subtypes and stages of melanocyte differentiation investigated in GSE45227 was calculated as follows: i) we selected the genes differentially expressed in at least one subtype; ii) for these genes we calculated the centroids (average expression) in each subtype of TCGA-SKCM and in each differentiation stage in GSE45227; iii) the Pearson's correlation coefficient between TCGA-SKCM and GSE45227 centroids was calculated and the results were visualized through a heatmap.

To evaluate the level of immune infiltration, stromal content and tumour purity we applied ESTIMATE (Yoshihara et al., 2013). This algorithm is based on single sample GSEA and uses an immune and a stromal signature to generate an immune and a stromal score. These two scores are then combined to obtain the estimate score from which tumour purity is derived. CIBERSORT (Newman et al., 2015) was applied to infer the immune cell composition of bulk tumours. CIBERSORT is a deconvolution method that quantifies the fraction of 22 immune cell subsets from gene expression data of bulk tumours. Samples with a significant deconvolution (CIBERSORT p-value < 0.05) were considered infiltrated by immune cells. This subset of samples was used to test the association between abundance of specific immune cells and consensus subtypes.

3.8.2 Proteomics

Differential protein expression analysis was performed using limma (Phipson et al., 2016). Correlation between matched reverse-phase protein array (RPPA) and RNA-Seq data was calculated using the Pearson's correlation coefficients. The distribution of the observed correlation coefficients was compared to a null distribution obtained by correlating proteins with random genes for 1000 iterations.

3.8.3 MicroRNAs

MicroRNA isoforms raw counts were preprocessed using the same voom/limma pipeline used for RNA-Seq data. Briefly, TMM was applied for between-samples normalization and normalized raw counts were transformed in log2-cpm. Differentially

expressed miRNAs were identified using limma after voom transformation. *In silico* prediction of miRNA targets was performed with six algorithms (DIANA MicroT-CDS (Reczko et al., 2012), microRNA.org database (Betel et al., 2008), mirDB (Wang, 2008), PITA (Kertesz et al., 2007), RNA22 (Miranda et al., 2006), and TargetScan v6.2 (Lewis et al., 2005)). Only targets predicted by at least two algorithms and negatively correlated with the miRNAs (Pearson's correlation coefficient < -0.2) were selected. The biological function of each miRNA was inferred from the pathways significantly enriched in the list of its targets. Significantly enriched pathways were identified using the gProfileR package (Reimand et al., 2016) with GO biological process terms, KEGG and REACTOME pathways. We then calculated their similarity in terms of shared pathways using Jaccard index. This similarity matrix was clustered using hierarchical clustering with Euclidean distance and Ward linkage to find groups of miRNAs involved in similar biological pathways.

3.8.4 DNA methylation

DNA methylation data were obtained from Illumina Infinium HumanMethylation450 Array platform. The methylation level of each CpG site is calculated from the fluorescent signals of the methylated and unmethylated alleles through the β-value that is the ratio of the methylated probe and the sum of methylated and unmethylated probe intensities. β-values range from 0 (totally unmethylated) to 1 (totally methylated). For statistical analyses β-values are converted to M-values that are the log2 ratio of the intensities of methylated probe versus unmethylated probe, while β-values are used for visualization purposes. Level 3 β-values were filtered removing CpG probes with any missing value, CpG probes mapping on X and Y chromosomes and CpG probes located 1 or 2 nucleotides from known single-nucleotide polymorphism with a minimum minor allele frequency of 0.05. The overall methylation level of each sample was calculated by averaging the β-values of all CpG probes. For differential methylation analysis β-values were converted to M-values and differentially methylated sites were identified using the DMRcate package (Peters et al., 2015). The methylation level of each gene was calculated by averaging the M-values of the CpG probes mapping on the promoter region, identified using the annotatr package (Cavalcante and Sartor, 2017). Correlation between gene-level methylation data and RNA-Seq data was calculated using the Pearson's correlation coefficient. The observed distribution of correlation coefficients was compared to a null distribution obtained by correlating methylation data with expression data from random genes for 1000 iterations. GSEA was performed on the list

of genes ranked according to their Pearson's correlation coefficient between matched methylation and expression data using the fgsea package (Sergushichev, 2016) with GO biological process gene sets and consensus subtypes gene sets (genes up-regulated in each consensus subtype with a fold change > 3 and an FDR < 0.05).

3.8.5 Somatic mutations

Genomic subtypes were defined according to the mutational status of BRAF, (N/H/K) RAS and NF1 genes. Patients with BRAF mutation at V600 or K601 aminoacid residues were assigned to the BRAF hotspot mutant subtype. Patients with mutations at residues G12, G13, Q61 in (N/H/K) RAS were assigned to the RAS hotspot mutant subtype. Non-synonymous mutations in NF1 gene defined the NF1 mutant subtype. Remaining patients were assigned to the triple WT subtype. DeconstructSigs package (Rosenthal et al., 2016) was used to perform the mutational signature analysis. This tool evaluates the contribution of 30 signatures reported in COSMIC (https://cancer.sanger.ac.uk/cosmic/signatures) to the mutational profile of each sample. The obtained signature scores were then analyzed in association with consensus subtypes. The maftools package (Mayakonda and Koeffler, 2016) was used to parse mutational maf files. The number of mutations for each patient was calculated excluding synonymous ones. To identify non-synonymous mutations associated to consensus subtypes we defined a binary mutational state for each gene (0 if WT; 1 if any mutation) before performing Fisher's exact test.

3.8.6 Somatic copy number aberrations

Tumour ploidy and purity estimated by ABSOLUTE algorithm (Carter et al., 2012) were downloaded from The Cancer Immunome Atlas database (https://tcia.at/) (Charoentong et al., 2017). Segmented copy number data were analyzed using GISTIC2.0 (Mermel et al., 2011) to identify significantly recurrent somatic copy number aberrations across the entire TCGA-SKCM cohort, independently of consensus subtypes. Significant focal copy number alterations were then tested for association with consensus subtypes. For this analysis thresholded copy number data were used, considering amplifications and deletions separately. For amplifications, a gene was assigned a value of 1 if amplified or 0 if the gene was WT. The same criteria were applied to deletions. The binary amplifications and deletions data were then used for the statistical analysis. GISTIC output was parsed using the maftools package (Mayakonda and Koeffler, 2016). The number of

amplifications and deletions per sample was calculated from the thresholded copy number data. Correlation between gene-level copy number and RNA-Seq data and GSEA were performed as described previously for methylation data.

4 RESULTS

4.1 Melanoma classifiers and their overlap

Nine gene expression-based melanoma classifications have been published and despite a concordance in the identification of two major groups defined by an invasive or a proliferative phenotype, discrepancies exist and a unified classification has not been derived yet. Before applying our bioinformatics approach to define a consensus classification, we evaluated the gene overlap of the nine gene expression signatures, reasoning that classifiers with a high number of shared genes would give concordant predictions. To measure the similarity of the selected signatures in terms of shared genes we calculated the Jaccard index between each pair-wise combination of signatures (36 combinations in total). Results showed a poor overlap between classifiers with a maximum level of similarity of 0.25 between Widmer and Sensi signatures (**Figure 4.1A**). Due to the wide range of the number of genes included in the signatures (from 55 genes of Rambow to 1495 genes of TCGA) we did not expected high Jaccard values for gene lists with very different size. However, even for signatures of comparable size the overlap was still poor.

We selected genes present in at least 5 signatures (> 50%) to identify a core set of recurring genes. No genes recurrent in all signatures were found. The most recurrent gene was MLANA, present in 8 out of 9 signatures, followed by DCT, GPR143, MITF and TYR, included in 7 signatures (**Figure 4.1B**). This core set of 30 genes was significantly enriched in targets of MITF according to IPA upstream regulators analysis ($p = 1.55e-37$), which are involved in the synthesis of melanin and in melanocyte development, suggesting that the most recurrent genes are representative of key biological processes in melanoma. Moreover, two different signatures with scarce overlap can represent different read-outs of the same biological process, thus we assessed the similarity among signatures in terms of biological function. The hypothesis is that reducing the heterogeneity among signatures shifting from a gene-level to a pathway-level the overlap of the signatures should increase. To this aim we performed over-representation analysis for each signature using IPA and we selected significantly enriched canonical pathways with an FDR < 0.05. We next evaluated the overlap of significantly enriched pathways between each pair of signatures (**Figure 4.1C**). Again, the number of significantly enriched pathways for each signature is highly variable, ranging from 1 for Tsoi and Tirosh to 167

for Verfaillie. The number of significantly enriched pathways is slightly correlated to the number of genes in each signature, despite not statistically significant (**Figure 4.1D**). Due to this heterogeneity a general increase in the Jaccard index values was not observed. Only four pair-wise combinations showed an increase of at least 0.1 when shifting from gene- to pathway-level (Jonsson-Dugo, Jonsson-TCGA, Dugo-Widmer and Sensi-Verfaillie). Four pathways were significantly enriched in at least 50% of signatures and included "*hepatic fibrosis/hepatic stellate cell activation*", "*WNT/β-catenin signaling*", "*melanocyte development and pigmentation signaling*" and "*eumelanin biosynthesis*" (**Figure 4.1E**).

Overall, these results showed that the nine melanoma classifiers were very heterogeneous in terms of number and type of genes included and this heterogeneity was mirrored at the pathway-level. However, it was possible to identify a core of recurrent genes and pathways that converged on MITF transcriptional regulation.

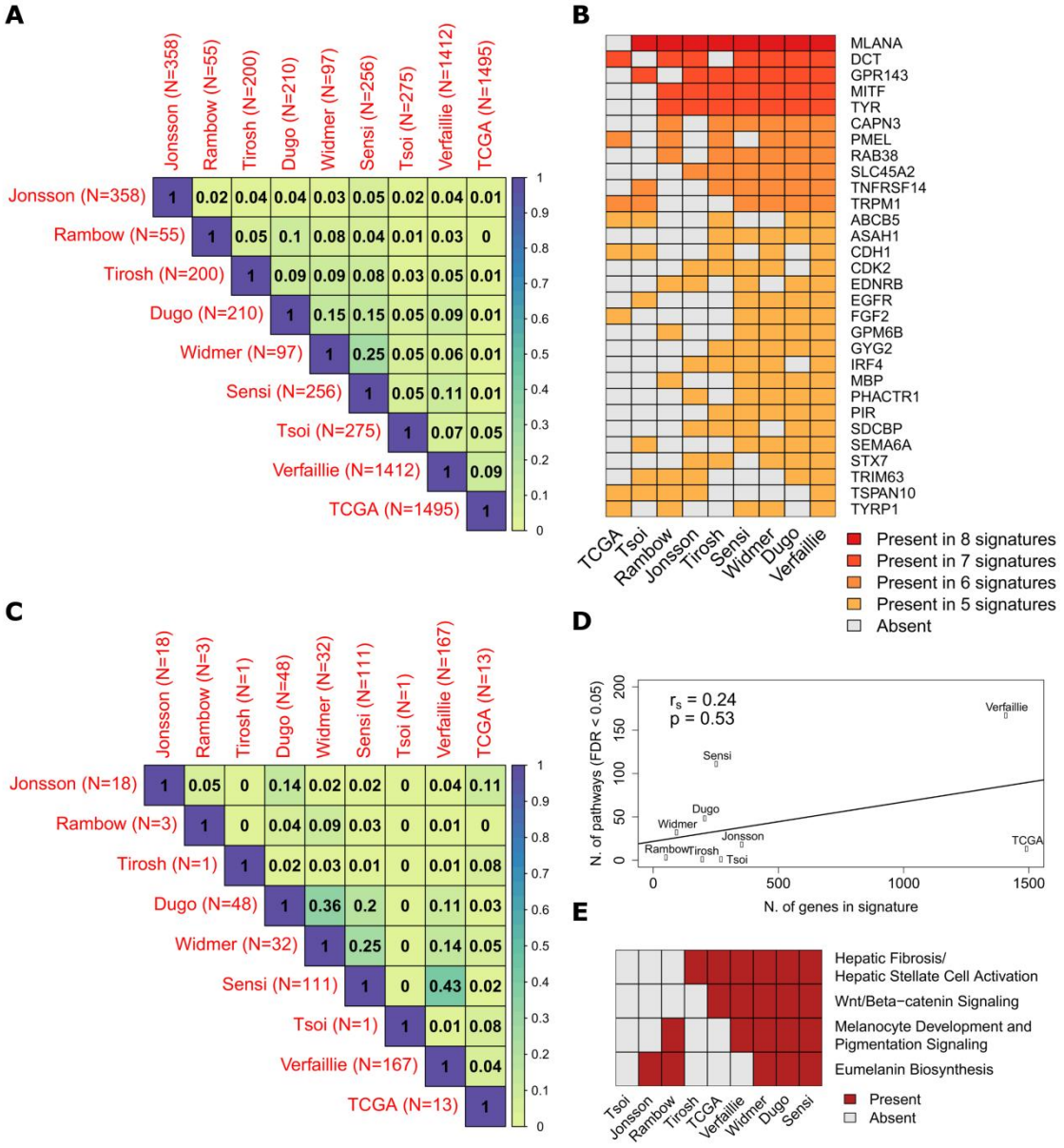

Figure 4.1. Similarity of published melanoma gene expression signatures. A) Heatmap of Jaccard index values between each pair of signatures. The color intensity is proportional to the magnitude of Jaccard index, reported in each matrix cell. N: number of genes in the signature. B) Heatmap showing the genes recurring in at least 5 signatures. C) Heatmap of Jaccard index calculated from the number of enriched pathways in each signature. N: number of genes in the signature. D) Spearman correlation between the number of genes and the number of significantly enriched pathways in each signature. E) Heatmap of enriched pathways in at least 5 signatures.

4.2 Identification of melanoma consensus molecular subtypes

To explore how the different classifiers divided melanoma tumours and to identify a robust consensus classification we applied an analytical workflow to two transcriptomic datasets. The first dataset included 471 cutaneous melanomas profiled using RNA-Seq by the TCGA consortium (TCGA-SKCM). The second dataset was assembled from 15 public gene expression studies of melanoma tumours analyzed using Illumina or Affymetrix microarray platforms, for a total of 897 samples. For this dataset, hereafter called MCSM

(Melanoma Clinical Samples from Microarrays), raw data from each study were appropriately pre-processed and combined at the gene level using ComBat to correct for batch effects and reduce technical variation among studies.

We assigned each sample of the two datasets to melanoma subtypes according to the nine classifiers using the NTP algorithm, a nearest neighbor-based class prediction method that provides the predicted class and a measure of significance of the assignment using only a list of signature genes. The percentage of samples assigned to a subtype by each classifier was comparable between TCGA-SKCM and MCSM datasets (**Figure 4.2**), confirming that both datasets are representative of melanoma phenotypic spectrum. Each signature was able to classify more than 80% of samples with the exception of Rambow and Tirosh classifiers that were characterized by a higher number of undetermined samples. The number of genes in Rambow classifier is very limited (n = 55) while Tirosh signature was derived from single-cell RNA-Seq data that are known to have a high level of technical noise in comparison to classical RNA-Seq or microarrays. These limitations could explain, at least in part, the reduced performance of the two classifiers.

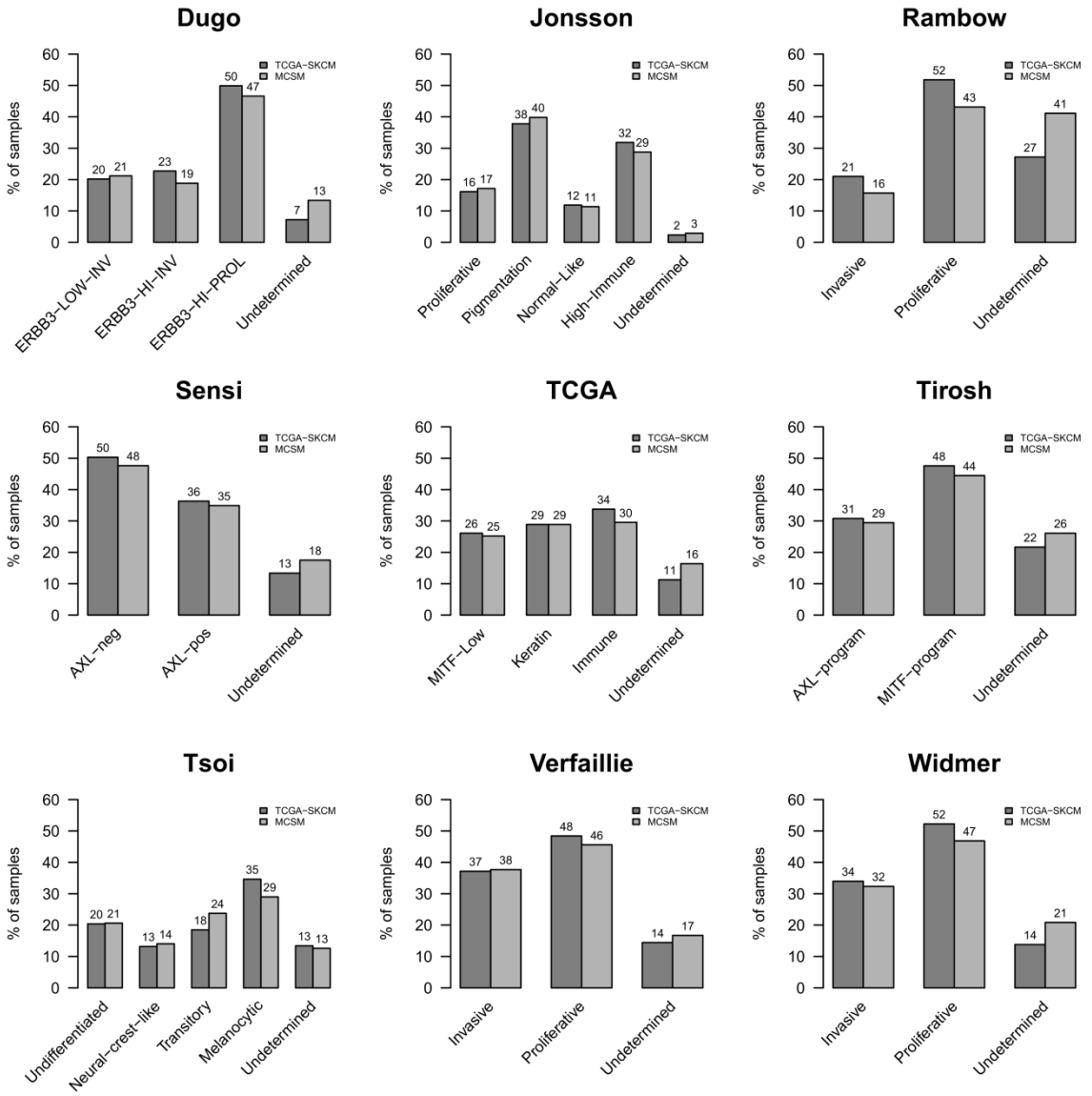

Figure 4.2. Distribution of subtype labels in TCGA-SKCM and MCSM datasets. Bar plots showing the percentage of samples assigned to each subtype according to the nine published signatures. For each classifier, a sample was significantly assigned to a subtype if the NTP FDR was less than 0.25. Samples exceeding this threshold were labeled "undetermined".

After NTP classification we obtained for each sample nine subtype labels. The classifications of TCGA-SKCM and MCSM samples were then combined in a matrix containing the nine subtype labels for all 1368 samples. In this way, the labels that were obtained from each dataset separately could be merged without introducing noise or artifacts due to the technological differences between the two datasets. To identify a consensus classification we calculated the similarity between samples according to their labels using the Hamming distance that corresponds to the number of different labels between two samples. Samples were then clustered according to this distance measure using consensus clustering in combination with the PAM algorithm and varying number of clusters K (K = 2 to K = 10). To estimate the optimal number of clusters we visually

examined for each K the consensus matrices, the cumulative distribution function (CDF) plot and the delta area under the CDF (ΔCDF) plot. The consensus matrix is a symmetric matrix, where both rows and columns represent samples and collects the proportion of times in which two samples are clustered together across the resampling runs (n = 1000). The CDF plot shows the CDF of the consensus matrix values at each K allowing visualizing at which K the CDF reaches a maximum. The ΔCDF plot shows the relative change in area under the CDF curve between K and K-1 to determine at which K there is no appreciable increase in the consensus. Evaluation of these plots pointed out that the optimal number of cluster was 5 as there was no noticeable increase in area under CDF function after K = 5 (**Figure 4.3B-C**) and the consensus matrix at K = 5 showed five well-separated blocks on the diagonal, even if clusters 4 and 5 showed a certain degree of overlap for some samples (**Figure 4.3A**). The same consensus analysis was repeated for TCGA-SKCM and MCSM independently (**Figure 4.4**). In both datasets we found that the optimal number of clusters was 5 reinforcing the reliability of the partitioning.

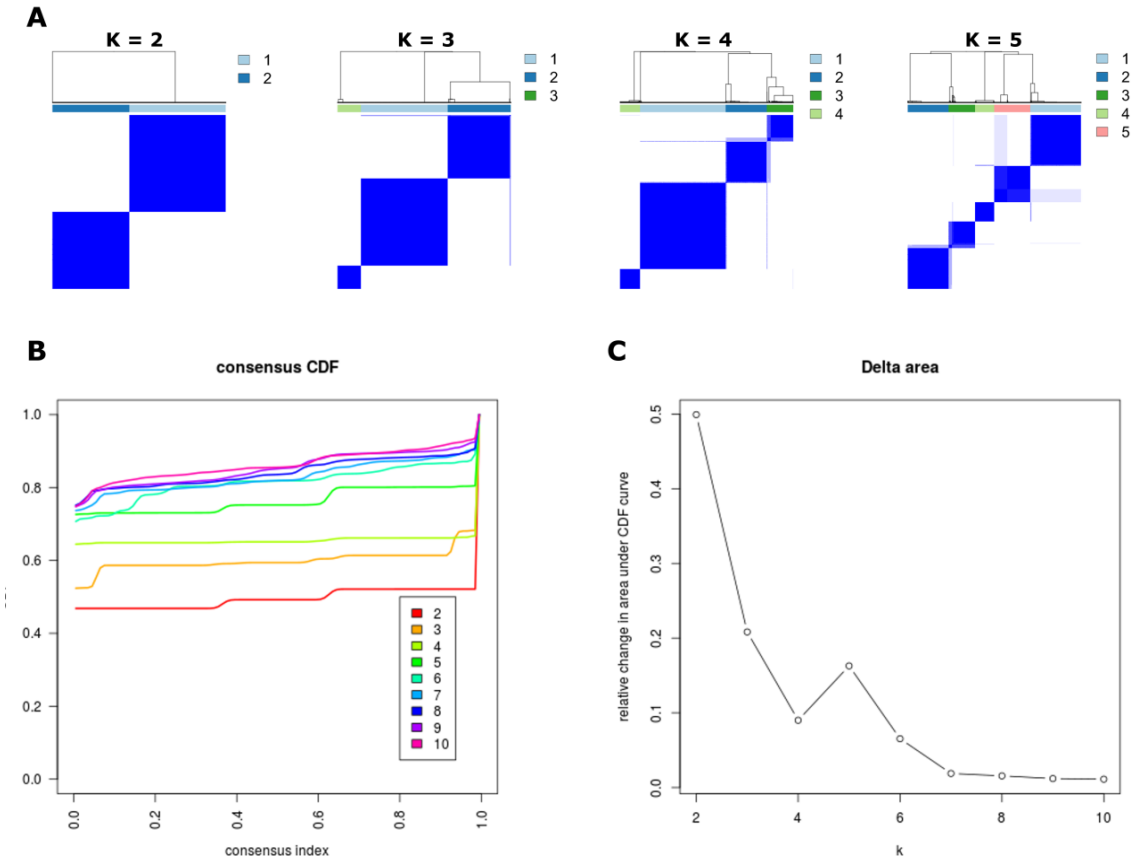

Figure 4.3. Consensus PAM clustering subtypes using TCGA-SKCM and MCSM combined labels. A) Consensus matrices for k = 2 to k = 5. Rows and columns represent samples and consensus values range from 0 (never clustered together) to 1 (always clustered together) marked by white to dark blue. B) CDF plot shows the cumulative distribution functions of the consensus matrix for each k. At k = 5 the CDF reaches a maximum with negligible improvement at k > 5. C) The relative change in area under CDF curve shows that as k is increasing beyond 5, there is a significant drop in the relative change in area under CDF curve, indicating an optimum at k = 5.

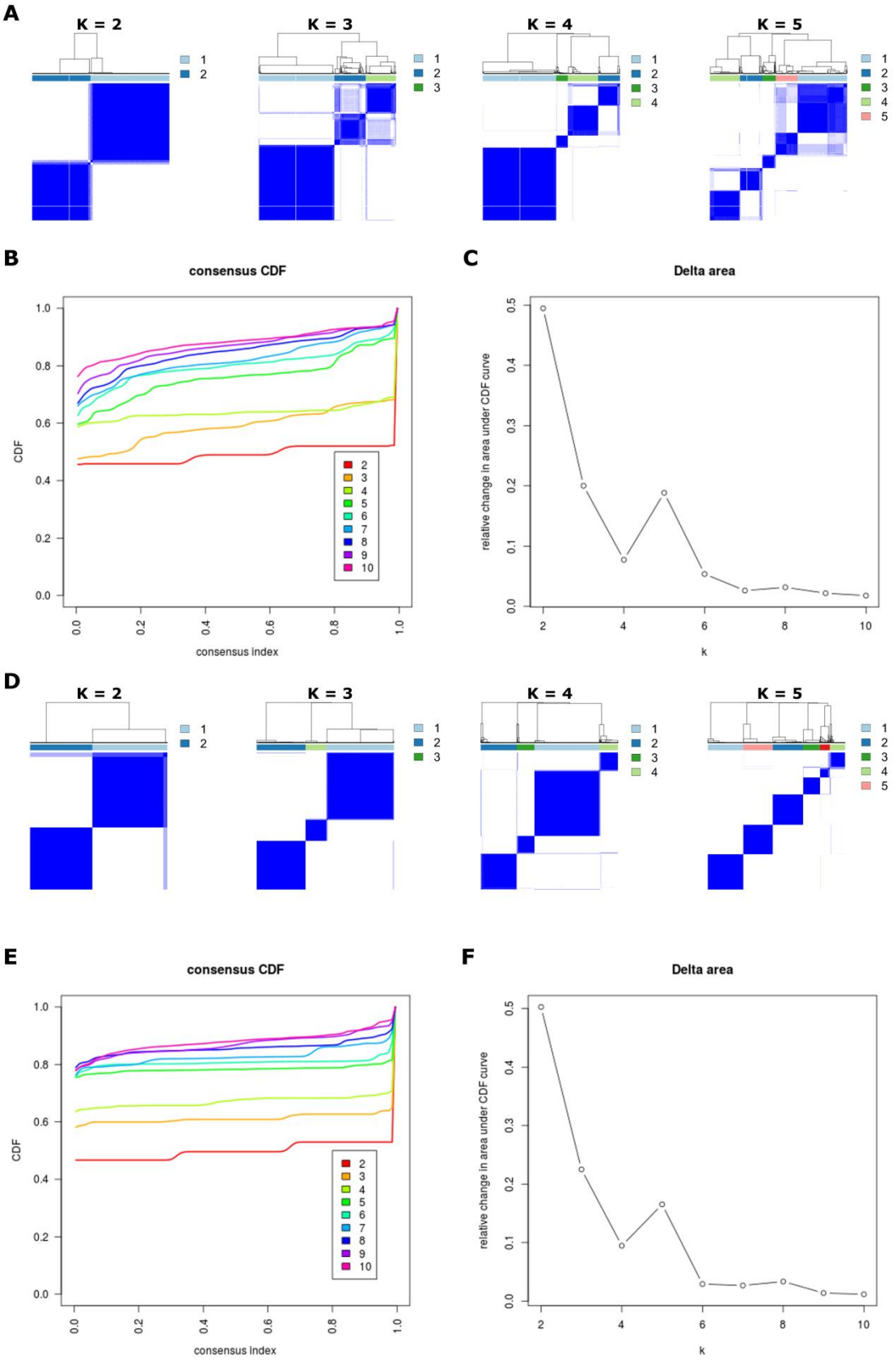

Figure 4.4. Consensus PAM clustering subtypes using TCGA-SKCM (A-C) and MCSM (D-F) separately. A, D) Consensus matrices for k = 2 to k = 5. Rows and columns represent samples and consensus values range from 0 (never clustered together) to 1 (always clustered together) marked by white to dark blue. B, E) CDF plot shows the cumulative distribution functions of the consensus matrix for each k. At k = 5 the CDF reaches a maximum with negligible improvement at k > 5. C, F) The relative change in area under CDF curve shows that as k is increasing beyond 5, there is a significant drop in the relative change in area under CDF curve, indicating an optimum at k = 5.

We then measured the overlap of consensus labels identified by PAM clustering with tumour labels from the nine classification systems. The overlap was visualized through the use of a heatmap (**Figure 4.5A**). We also statistically assessed the overlap using Jaccard index and hypergeometric test. Significant overlaps (Jaccard index ≥ 0.25 and hypergeometric FDR < 0.05) were displayed as a network (**Figure 4.5B**). CS2 and CS3 were representative of invasive samples with CS2 being enriched of "ERBB3-LOW-Invasive", "Undifferentiated", "Immune" and "High-immune" subtypes and CS3 of "ERBB3-HI-Invasive", "Neural crest-like", "Proliferative" from Jonsson and "MITF-low". Interestingly, CS2 was also enriched of "Undetermined" samples from Rambow signature. This observation suggested that this signature, containing only 55 genes, did not recognize these samples as classical invasive, supporting the splitting of the invasive subtype in two sub-groups. In the same way, CS1 and CS4 both included proliferative samples but CS1 was enriched of "ERBB3-HI-Proliferative","MITF-program","Keratin" and "Pigmentation" while CS4 consisted of "Transitory" samples. Finally, CS5 was enriched in samples classified as "Undetermined" by the majority of signatures, indicating the existence of a fifth group that could not be discriminated by the actual classifiers.

Overall, these results showed that melanoma tumours can be classified in five robust subtypes, with four of them corresponding to previously described biological entities. In addition, a new subtype that comprises samples classified as "Undetermined" by the nine classification systems was discovered. This last subtype could represent a new biological subset of melanomas not previously identified.

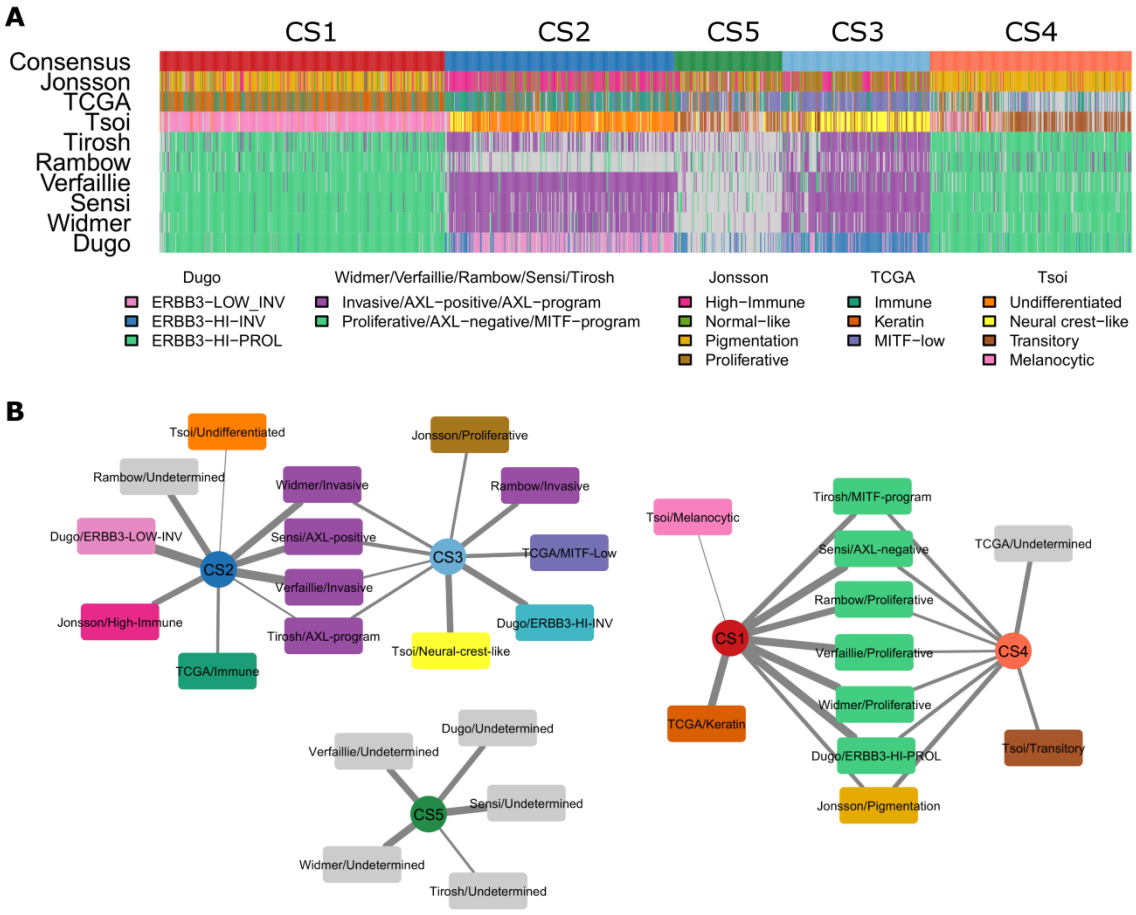

Figure 4.5. Comparison of consensus subtypes and the original classification systems. A) Overlap between consensus subtypes (CS1-CS5) and labels assigned by the nine melanoma gene classifiers for the combined TCGA-SKCM and MCSM datasets. B) Association of consensus subtypes with subtype labels from the 9 classification systems. Each node is a subtype label (circular node are consensus subtypes) colored according to the color scheme in A). Edge width is proportional to the overlap between labels measured using Jaccard index. Only significant overlaps (hypergeometric FDR < 0.05) are shown.

4.3 Differential pathway activation of consensus molecular subtypes

4.3.1 Transcriptomics

To assess transcriptional differences among subtypes on a pathway-based level we applied GSVA to TCGA-SKCM dataset, comparing each subtype with the others. We first focused on the previously reported invasive/proliferative gene signatures (**Figure 4.6A**). As expected, CS2 and CS3 subtypes showed a strong up-regulation of invasive signatures and were depleted for proliferative ones. Vice versa, CS4 and CS1 groups showed an opposite enrichment pattern. These results confirmed that these four subtypes were a further subdivision of the two major invasive and proliferative phenotypes, as already demonstrated in previous works (Dugo et al., 2015; Tsoi et al., 2018). Interestingly, the CS5 subtypes showed a negative enrichment of both invasive and proliferative signatures. Investigation of the expression pattern of single genes relevant for melanoma biology and classification confirmed this finding. MITF and its target genes

MLANA and PMEL were expressed at lower levels in CS2 and CS3 subtypes while they were up-regulated in CS4 and CS1 (**Figure 4.6B**). The opposite pattern was observed for the receptor tyrosine kinases AXL and EGFR. Also SOX10, a major regulator of neural crest development, was down-regulated in the CS2 subtype in comparison to the CS4 and CS1 proliferative groups. For all these genes CS5 subtype showed an intermediate level of expression. These observations indicated that CS5 subtype could be a turning point between invasive and proliferative phenotypes, preceding the CS4 subtype along the melanocyte differentiation process.

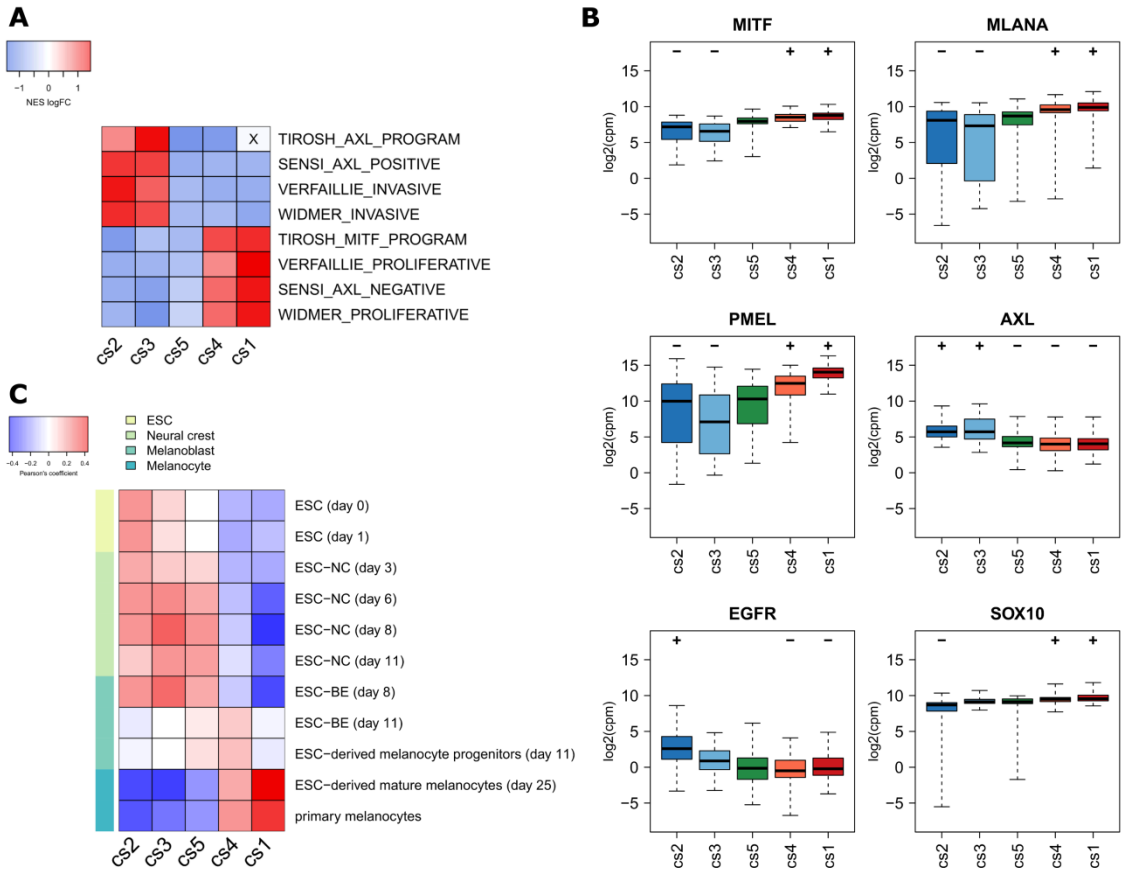

Figure 4.6. Phenotype switching and differentiation features of consensus subtypes. A) GSVA analysis for invasive and proliferative gene sets. The heatmap shows the log fold change of the normalized enrichment scores obtained comparing each subtypes versus the others. An "X" denotes that the FDR associated to the fold change is not significant (FDR ≥ 0.05). B) Box plots of expression levels for selected relevant genes involved in phenotype switching. A "+" or a "-" indicate that the gene is significantly up- or down-regulated, respectively. C) Heatmap of the Pearson's correlation coefficient between the centroids of consensus subtypes and of experimental conditions of the melanocyte differentiation in vitro model from GSE45227.

To confirm this hypothesis we compared subtypes with a gene expression dataset recapitulating an *in vitro* model of human melanocyte differentiation (GSE45227, (Mica et al., 2013). This dataset consisted of a time-course experiment of human embryonic stem cells (ESC) sequentially treated to induce differentiation to neural crest cells, melanoblasts and mature melanocytes. Tsoi et al (Tsoi et al., 2018) already compared

their melanoma subtypes with the melanocyte differentiation stages using this same dataset. However, they considered for the comparison only selected time points representative of the major differentiation states but excluded all other time points that could represent intermediate states. Here we performed a more detailed analysis comparing the five consensus subtypes with all time points available. To this aim we calculated for the five consensus subtypes identified in TCGA-SKCM dataset the centroid, i. e. the mean expression profile of samples within the same subtype. For this calculation we considered only the genes that were significantly differentially expressed in at least one subtype (ANOVA FDR < 0.05). Using the same subset of genes we next calculated the centroids of each time point in the differentiation dataset. Finally, we calculated the correlation between the centroids of the two datasets. The CS2 subtype correlated well with ESC stage, while CS1 resembled the mature melanocytes phenotype, with the CS3 and CS4 representing the intermediate stages of neural crest and melanoblast, respectively. As hypothesized, the CS5 located in between the CS3 and CS4 subtypes, representing a late neural crest/early melanoblast stage (**Figure 4.6C**).

To gain a broader overview of the biology underlying each specific subtype we repeated GSVA for each group compared with the remaining four using the Gene Ontology (GO) terms of the biological process domain. We identified several significantly enriched pathways for each subtype, supporting subtype-specific transcriptional alterations. The top-100 GO terms significantly enriched in each subtype (FDR < 0.05) were visualized as a network to identify clusters of overlapping gene sets attributable to the same biological process (**Figure 4.7**). In CS2 we observed an enrichment of genes involved in several signaling pathways such as RTK, PIK3CA, MAPK and JAK/STAT signaling and in immune system. Angiogenesis, cellular migration, extracellular matrix remodeling and neural crest differentiation were enriched in the CS3 subtype. In the two proliferative subtypes (CS4 and CS1) we found enrichment of gene sets related to pigment cell differentiation and pigment biosynthesis but CS4 was specifically enriched of genes involved in transcription, mRNA processing and splicing, DNA repair and DNA methylation. On the other side CS1 was preferentially enriched of metabolic gene sets.

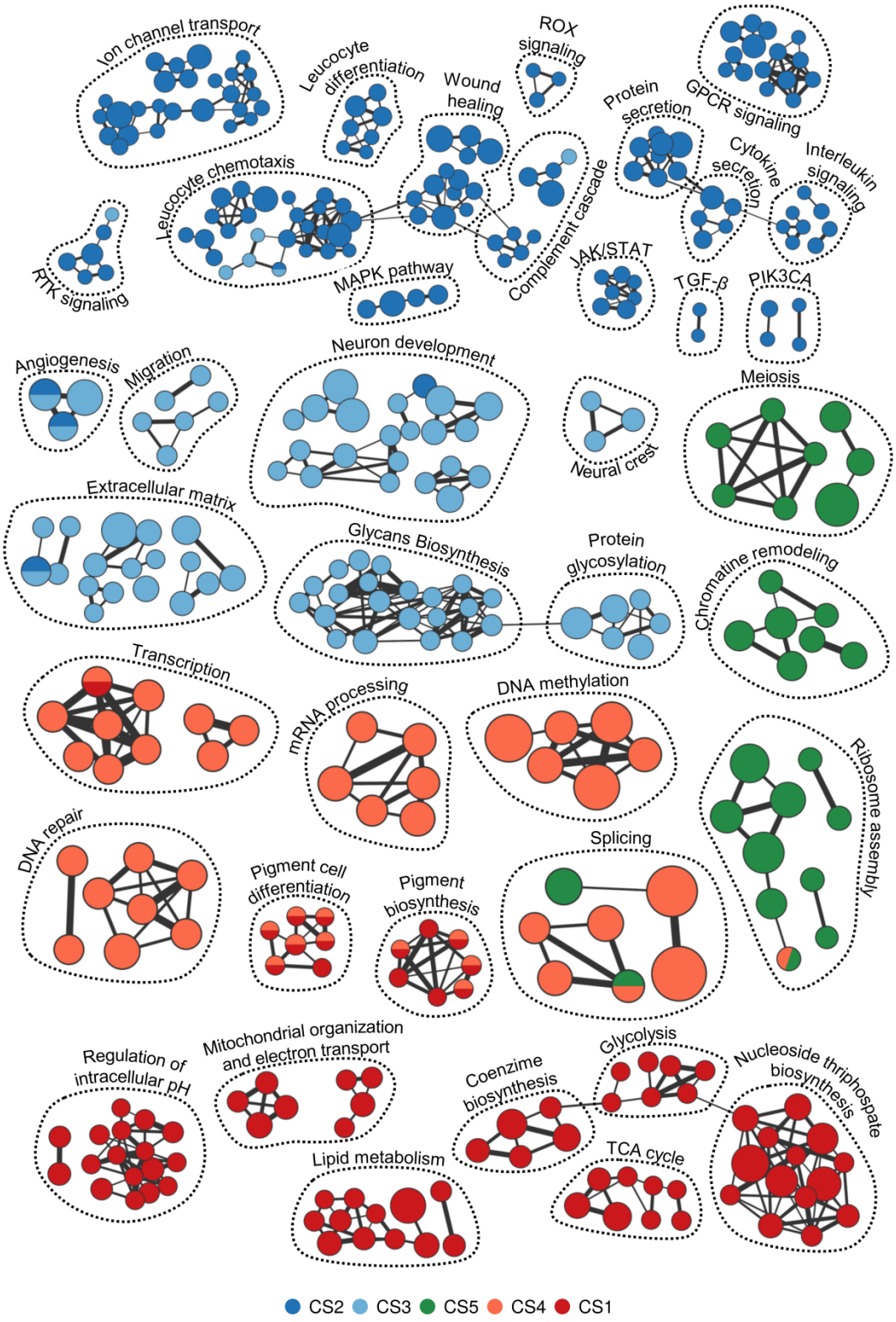

Figure 4.7. Subtype-enriched pathways. Network showing the top-10% GO biological process terms in each subtype. Each node is a GO term and node size is proportional to the number of its genes. Nodes with genes in common are connected by an edge whose width is proportional to the number of shared genes. Nodes are colored according to the subtype in which they are enriched; multiple colors indicate that the gene set is enriched in more than one subtype. Cluster of nodes sharing several genes were manually grouped in biological macro-categories.

The newly discovered CS5 subtype showed up-regulation of genes involved in ribosome assembly, chromatin remodeling and meiosis. Many regulators of meiosis are

expressed in germ cells and are silenced in healthy somatic tissues but they can be aberrantly expressed in cancerous tissues inducing chromosomal instability (CIN). These genes are referred as cancer-testis antigens (CTAs) and they are frequently expressed in melanoma. Given these premises we hypothesized that the CS5 subtype could be characterized by over-expression of CTAs and high CIN. Our hypothesis was confirmed by GSEA analysis that highlighted the strong positive enrichment of CTAs and of a CIN signature (Carter et al., 2006) in the genes up-regulated in this subtype (**Figure 4.8A-B**).

Overall, these results confirmed that the consensus subtypes recapitulated the differentiation stages from ESC to fully differentiated melanocytes and each subtype was characterized by the activation of specific pathways. To give the consensus subtypes more meaningful names we adopted the nomenclature already used in the work of Tsoi et al except for CS4 and the newly discovered CS5 (**Table 4.1**). Having shown that CS5 was the true intermediate state between the two invasive and proliferative subtypes, and it was characterized by high CIN, we named the CS5 subtype as "Transitory/CIN". The CS4 subtype that was the original transitory subtype from Tsoi et al (Tsoi et al., 2018) was renamed "Melanoblast-like", as our analyses showed that it corresponded to cells in the state of melanoblast along the differentiation axis.

Table 4.1. Nomenclature of melanoma consensus subtypes.

Consensus subtype	Name
CS2	Undifferentiated
CS3	Neural crest-like
CS5	Transitory/CIN
CS4	Melanoblast-like
CS1	Melanocytic

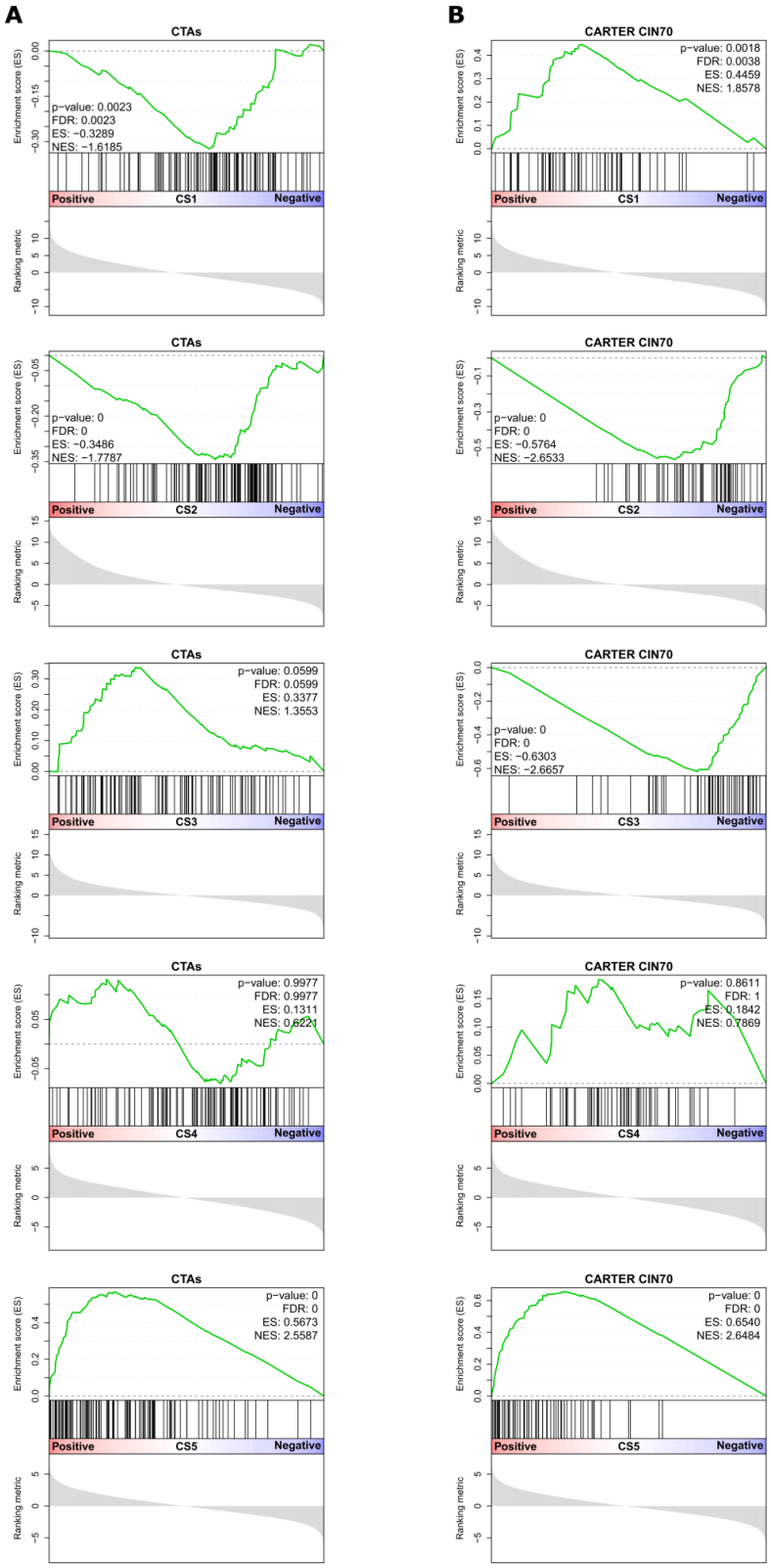

Figure 4.8. Enrichment of CTAs and CIN signatures in consensus subtypes The plots show the enrichment of CTA (A) and CIN (B) gene sets in each subtype. The grey bar plot at the bottom of each plot shows the genes ranked from the most up-regulated (right) to the most down-regulated (left) in each subtype. The vertical black bars show where the genes of the gene sets fall along the ranked gene list, together with the running enrichment score (green line). For each analysis the enrichment score (ES), normalized enrichment score (NES), p-value and FDR are reported.

4.3.2 Correlation between mRNA and protein levels

Reverse-phase protein array (RPPA) is a high-throughput proteomic technology that uses antibody binding to quantify protein expression. For a subset of 353 samples of TCGA-SKCM cohort both RNA-Seq and RPPA data were available. RPPA expression measurements were available for 156 total proteins and all of them had matched mRNA data. We first assessed to what extent protein expression levels correlated with their mRNA expression. Overall, we found that Pearson's correlation coefficients for matched proteins and genes were shifted toward positive values (median = 0.24) compared to the null distribution of the correlation coefficients (median = 0) calculated using random matching of protein and RNA-Seq data for 1000 iterations (**Figure 4.9A**). This observation indicated that protein and corresponding mRNA levels were generally well correlated supporting the functional effect of the alterations observed at the gene expression level.

We next investigated the association between protein expression and consensus subtypes. For each subtype compared with all other four we found several proteins significantly differentially expressed (**Figure 4.9B**).

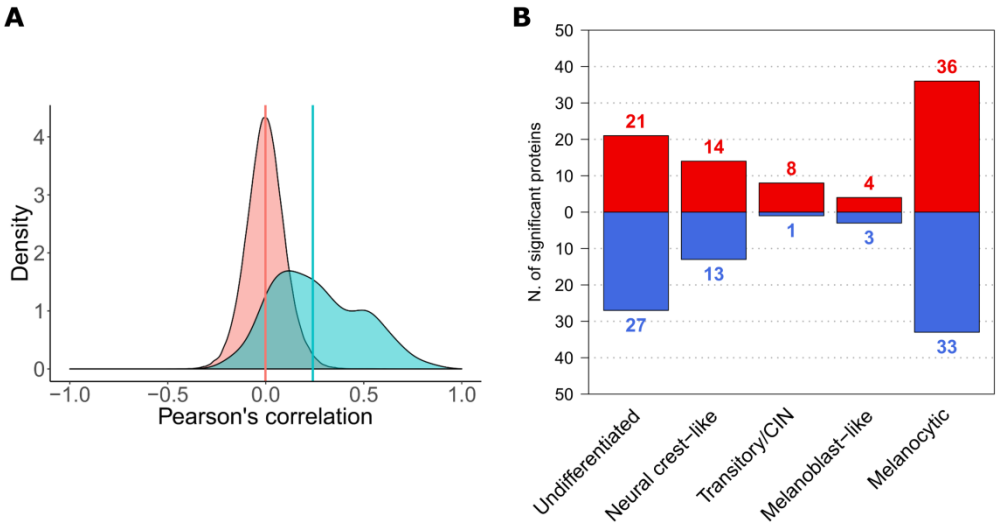

Figure 4.9. Analysis of TCGA-SKCM proteomic data. A) Distribution of the actual Pearson's correlation coefficients between each protein and its corresponding gene (green curve) compared to a random null distribution calculated by permuting genes 1000 times (red curve). Vertical lines represent the medians of the distributions. B) Number of significantly up-regulated (red) and down-regulated (blue) proteins (FDR < 0.05) for each subtype versus the others.

The undifferentiated subtype showed higher expression of Axl and reduced expression of Her3 as already observed at the mRNA level (**Figure 4.10**). Among the most up-regulated proteins in undifferentiated samples, we found markers of invasiveness such as Annexin-I, Annexin-VII and Caveolin-1. In addition, Syk and Lck, two proteins representative of T- and B-cell signaling, were highly up-regulated, confirming the high

prevalence of immune infiltrated tumours in this subtype. The neural crest-like subtype showed the same modulation pattern of the molecules described above for the undifferentiated group, confirming the invasive phenotype of these two subtypes. In addition, neural crest-like tumours showed up-regulation of additional markers of angiogenesis (IGFBP2) and invasiveness such as fibronectin, the integrin CD49b and PKC-alpha, in concert with loss of E-cadherin expression. We also found up-regulation of asparagine synthetase (ASNS) in this specific subtype. The two proliferative subtypes showed an opposite expression patterns for the proteins described above, more markedly in the melanocytic subtype. Compared to melanoblast-like tumours, the melanocytic subtype showed a higher expression of mTOR, Akt, E-cadherin and proliferation markers such as PCNA, Cyclin-E1 and Cyclin-B1. The melanoblast-like subtype showed specific up-regulation of STAT5-alpha, eEF2K and Her2. Finally the transitory/CIN subtype showed up-regulation of several DNA repair proteins such as MSH6, RBM15, GAB2 and Bim.

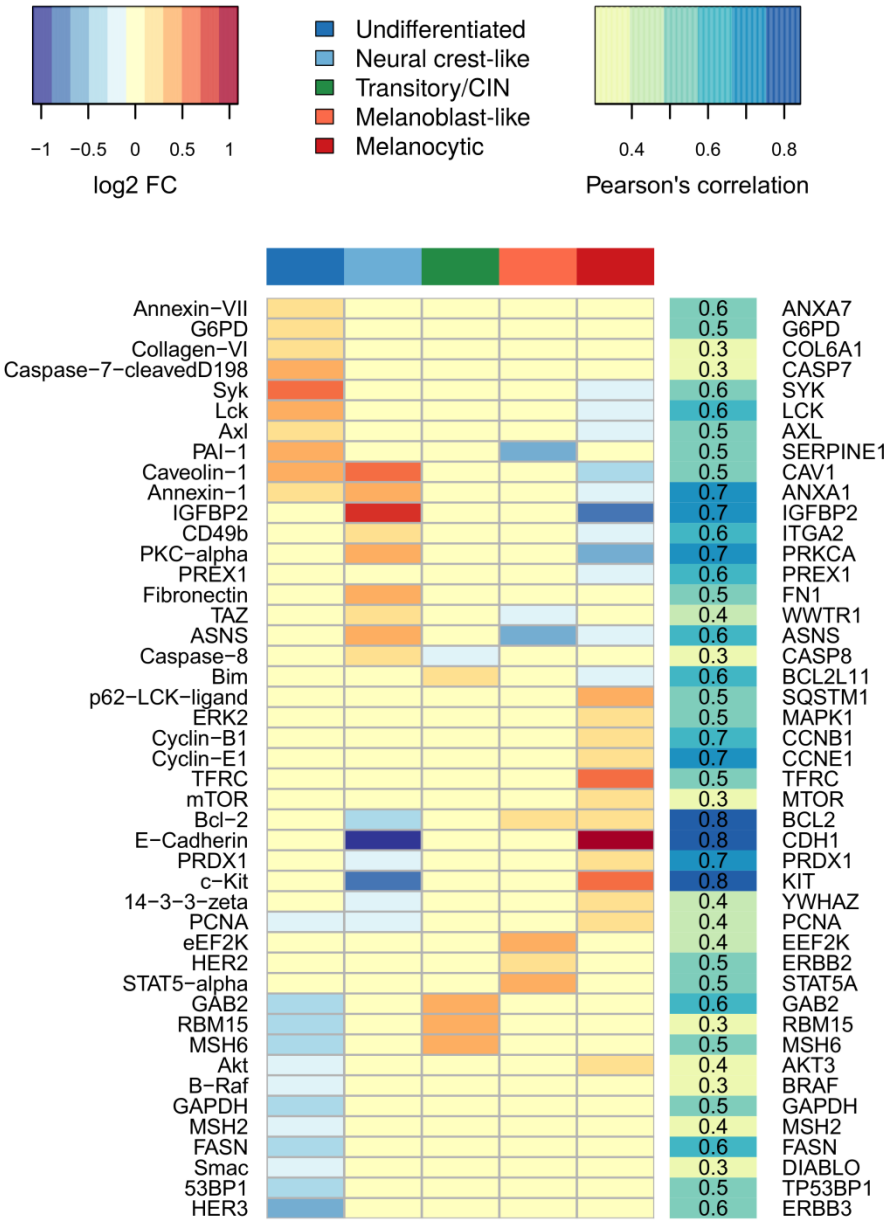

Figure 4.10. Heatmap of log fold changes for selected significant proteins. Yellow cells represent proteins not differentially expressed in the given subtype. The heatmap on the right shows the Pearson's correlation coefficient between the protein and the respective gene.

4.4 Microenvironment of consensus subtypes

The enrichment of immune-related gene sets in the undifferentiated subtypes made us hypothesize that consensus subtypes could be characterized by different tumour microenvironments. We applied ESTIMATE, a method that uses specific stromal and immune signatures to infer from gene expression data the fraction of stromal and immune cells in a bulk tumour sample. Stromal and immune scores are then combined to finally derive the sample tumour purity. Undifferentiated samples showed the highest immune score compared to all other (**Figure 4.11A**). The stromal score was highest in the undifferentiated subtypes but also neural crest showed a higher enrichment of stroma

compared to transitory/CIN, melanoblast-like and melanocytic (**Figure 4.11B**). Consequently, undifferentiated and neural crest-like samples were characterized by lower tumour purity (**Figure 4.11C**). To deeper investigate the microenvironment composition of consensus subtypes we performed GSVA using a set of cell-type specific signatures derived from a single-cell RNA-Seq study of melanoma samples (Tirosh et al., 2016). The results confirmed that undifferentiated samples were highly enriched in stromal and immune cells, including macrophages, B cells, endothelial cells and fibroblasts (**Figure 4.11D**). The neural crest-like subtype showed a positive enrichment of fibroblasts and endothelial cells, even if to a less extent, while it did not show enrichment of B cells. Oppositely, the two proliferative subtypes were enriched of melanocyte markers and showed a negative enrichment of stromal populations. The transitory/CIN subtype was not infiltrated by immune cells but was also negatively enriched of melanocyte markers, reinforcing the hypothesis that this subtype consisted of tumour cells in an intermediate state between invasive and proliferative phenotypes.

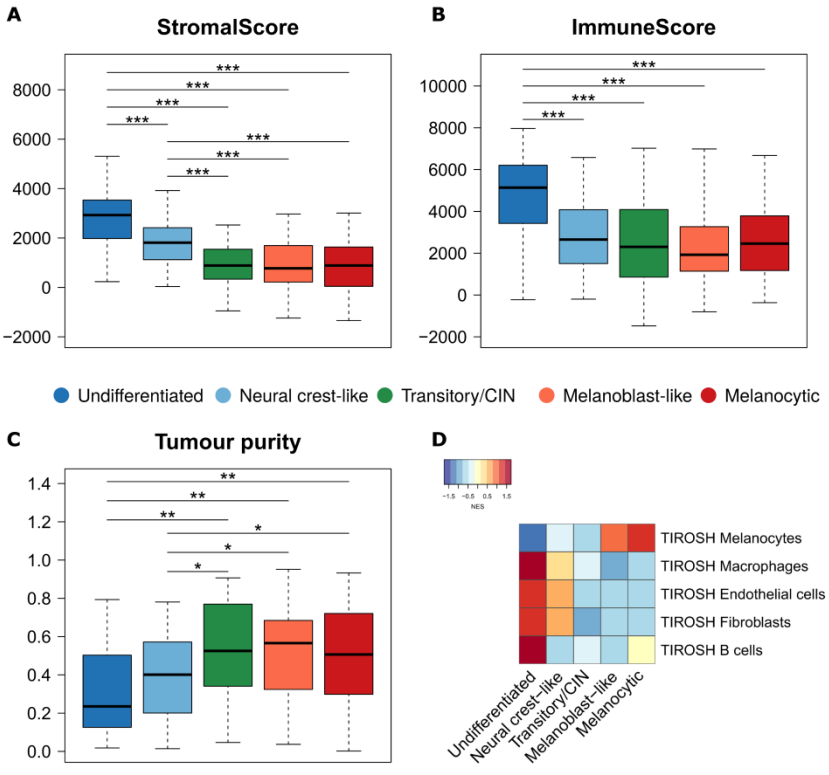

Figure 4.11. Microenvironment contribution to the gene expression of consensus subtypes. A-B) Box plot of the stromal and immune score calculated from the expression of specific markers using the ESTIMATE package. C) Tumour purity inferred from the combination of the stromal and immune score. P-values were obtained with ANOVA followed by Tukey's post hoc test. ***: $p < 0.001$; **: $p < 0.01$; *: $p < 0.05$. D) Heatmap of the average NES calculated using GSVA for cell-specific markers in each subtype.

To quantify the immune cellular composition of consensus subtypes we next applied CIBERSORT, a gene expression deconvolution method that can estimate the

relative proportion of 22 immune cell populations in bulk tumour samples. In addition to the estimated cell fractions, CIBERSORT returns a p-value of the deconvolution. This p-value tests the null hypothesis that none of the immune cells assessed by the algorithm are present in a given sample. Thus, samples with a significant deconvolution ($p < 0.05$) were considered "immune" while remaining samples were classified "non-immune". Association with consensus subtypes revealed that the majority of immune samples belonged to the undifferentiated subtype ($\chi 2$ p-value = 5.6e-22) (**Figure 4.12A**). Considering only samples with a significant CIBERSORT p-value, we evaluated the differential abundance of immune cell populations across consensus subtypes. Among the 22 subpopulations tested, 3 were significantly (ANOVA $p < 0.05$) differentially represented in the different subtypes (**Figure 4.12B-D**). The most significant difference in subpopulation content was observed between undifferentiated and neural crest-like subtype with macrophages M0 abundance and CD8+ T cells being lower and higher in the first subtype, respectively. Other significant differences were in both M0 and M2 macrophages among undifferentiated and melanocytic subtype.

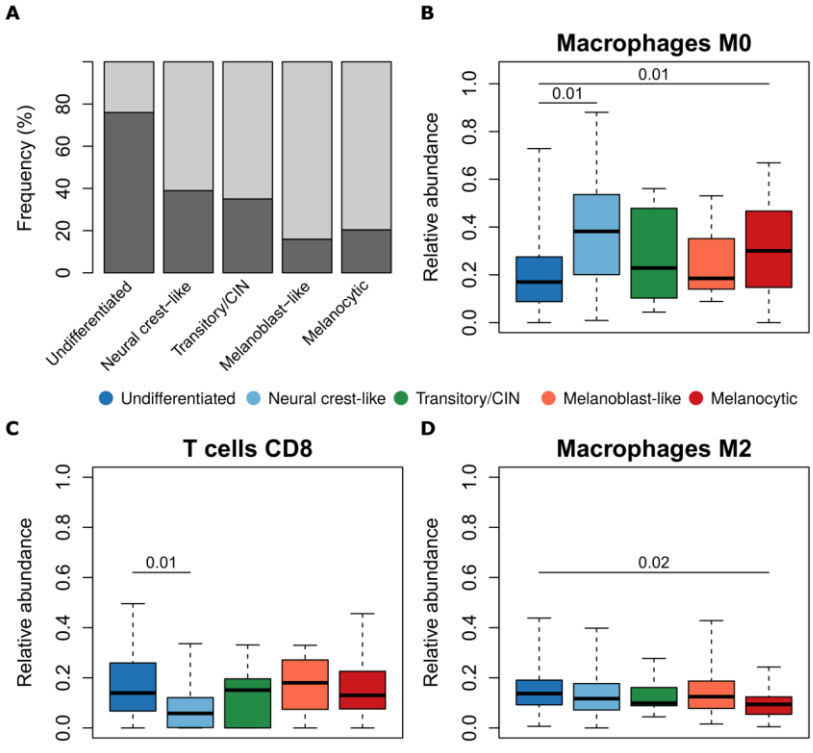

Figure 4.12. Association between immune infiltration status defined by CIBERSORT and consensus subtypes. The percentage of immune samples (CIBERSORT p-value < 0.05, dark grey) and non-immune samples (CIBERSORT p-value ≥ 0.05, light grey) are presented for each subtype. B-D) Box plots of CIBERSORT-inferred cell proportions for the cell subpopulations with an ANOVA p-value < 0.05. Reported p-values for pair-wise comparisons were calculated using Tukey's post-hoc test.

4.5 Prognostic impact of consensus subtypes

Identification of prognostic molecular biomarkers to improve the stratification of patients in different risk categories is a hot topic of research in oncology. In melanoma, patients with primary tumours can reach high rate of survival depending on stage at diagnosis while survival of patients with metastatic disease is poor. However, long survivors among patients with advanced disease can be observed, indicating that clinically approved prognostic biomarkers alone do not explain the observed variability in outcome. Here we assessed the prognostic impact of consensus subtypes in relation to overall survival for 161 TCGA-SKCM patients with stage III-IV metastatic disease. According to univariate Cox proportional hazards analysis, patients within the melanocytic and transitory/CIN subtype had an increased risk of death compared to all other groups (**Figure 4.13A** and **Table 4.2**). Neural crest-like patients had the best prognosis compared to all other subtypes. Based on previous findings reporting a protective role of immune infiltration for melanoma patients we categorized TCGA-SKCM samples in "Immune" or "Non-immune" according to the significance ($p < 0.05$) of the CIBERSORT p-value. Patients belonging to the immune class had significantly higher survival rates compared to the non-immune class (**Table 4.2** and **Figure 4.13B**). We next performed a multivariate Cox regression analysis to simultaneously assess the influence of consensus subtypes and immune status on overall survival. The two variables maintained their association with overall survival suggesting that consensus subtype and immune status jointly impact on survival (**Table 4.2**).

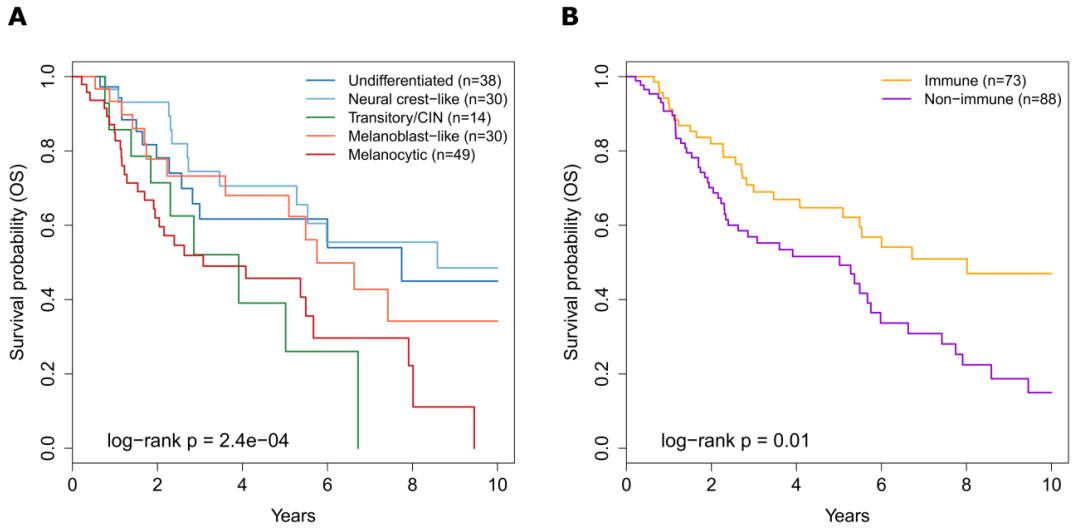

Figure 4.13. Prognostic impact of consensus subtypes and immune infiltration in stage III-IV metastatic melanoma of TCGA-SKCM dataset. A-B) Kaplan-Meier curves based on consensus subtypes (A) and immune status defined by CIBERSORT p-value (B). P-values were obtained through the log-rank test.

Table 4.2. Univariate and multivariate Cox regression analysis for consensus subtypes and immune status in metastatic patients of TCGA-SKCM dataset. Significant p-values are highlighted in bold. N: total number of samples; HR: hazard ratio; CI: confidence interval.

		Univariate analysis			Multivariate analysis		
Variable	**Events/N**	**HR**	**95% CI**	**p**	**HR**	**95% CI**	**p-value**
Consensus subtype							
Melanocytic	29/49	1	Ref	Ref	1	Ref	Ref
Melanoblast-like	14/30	0.52	0.27-1.01	0.053	0.59	0.30-1.17	0.130
Transitory/CIN	9/14	1.02	0.48-2.17	0.951	1.06	0.49-2.25	0.889
Neural crest-like	12/30	0.36	0.18-0.71	**0.003**	0.45	0.22-0.89	**0.023**
Undifferentiated	14/38	0.47	0.25-0.91	**0.026**	0.53	0.25-1.10	0.087
Immune status							
Non-immune	49/88	1	Ref	Ref	1	Ref	Ref
Immune	29/73	0.53	0.33-0.85	**0.009**	0.53	0.31-0.91	**0.020**

4.6 Identification of consensus subtypes in melanoma cell lines

Tumour-derived cell lines are the most common models used in cancer research to study the molecular mechanisms of cancer and response to drugs. Understanding how closely cell lines resemble *in vivo* tumours is crucial to evaluate the clinical significance of discoveries made using *in vitro* models. To this aim we applied the same unsupervised consensus PAM clustering approach used to identify the consensus subtypes in bulk melanoma tumours to a large gene expression dataset of melanoma cell lines. We opted to apply the unsupervised PAM approach, instead of deriving a predictor of the consensus subtypes from bulk tumour data, to avoid possible confounding in gene expression signals related to tumour microenvironment components.

As for bulk tumours, we first classified each cell line using the nine published signatures and then we applied consensus PAM clustering on the matrix of subtype labels. Inspection of the consensus matrices, area under the CDF curve and ΔCDF plot showed that the optimal number of clusters was 5 (**Figure 4.14**). This was in agreement with what was observed on TCGA-SKCM and MCSM datasets.

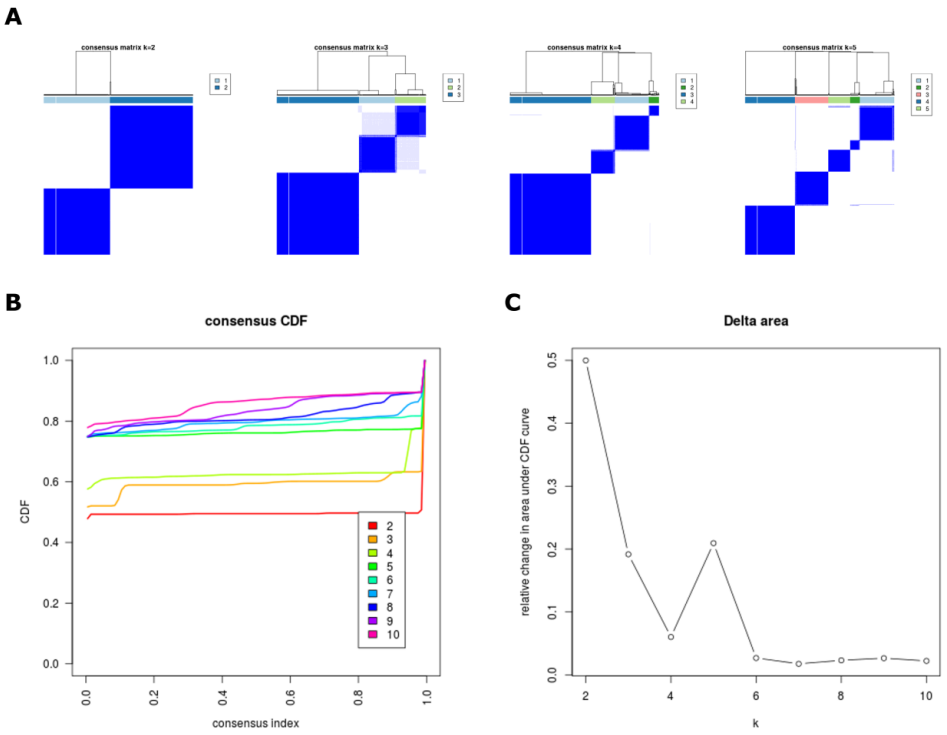

Figure 4.14. Identification of consensus subtypes through consensus PAM clustering in melanoma cell lines. A) Consensus matrices for k = 2 to k = 5. Rows and columns represent samples and consensus values range from 0 (never clustered together) to 1 (always clustered together) marked by white to dark blue. B) CDF plot shows the cumulative distribution functions of the consensus matrix for each k. At k = 5 the CDF reaches a maximum with negligible improvement at k > 5. C) The relative change in area under CDF curve shows that as k is increasing beyond 5, there is a significant drop in the relative change in area under CDF curve, indicating an optimum at k = 5.

We next verified whether the 5 subtypes identified in cell lines corresponded to the consensus subtypes from clinical specimens. The same pattern of overlap between the consensus classification and the labels from the nine published classifiers was observed (**Figure 4.15A**). CS2 included the subset of invasive cell lines classified as ERBB3-LOW-Invasive or undifferentiated by Dugo and Tsoi classifiers, respectively. The CS2 subtype was also enriched of cell lines classified as High-immune and Immune by Jonsson and TCGA. This suggested that the expression of immune-related genes was a tumour-intrinsic trait and it was not uniquely related to the infiltration of immune cells in bulk tumours. The second subset of invasive cell lines was enriched in the CS3 subtype. The CS4 and CS5 groups were enriched of transitory proliferative and keratin/melanocytic cell lines, respectively. Finally, the newly identified CS5 was present also in cell lines and grouped the majority of samples that were classified as undetermined by the nine predictors. As observed for clinical tumours, the 5 consensus subtypes identified *in vitro* recapitulated the different stages of melanocytes differentiation (**Figure 4.15B**). Finally, consensus subtypes signatures previously identified in TCGA-SKCM dataset were tested

for enrichment in the 5 five *in vitro* consensus subtypes. We observed that each signature was positively enriched in its correspondent *in vitro* subtype, with the exception of the melanocytic signature that was enriched in both CS1 and CS5 (**Figure 4.15C**). Collectively, these results showed that consensus subtypes and their expression patterns were consistent between bulk tumours and *in vitro* cell lines.

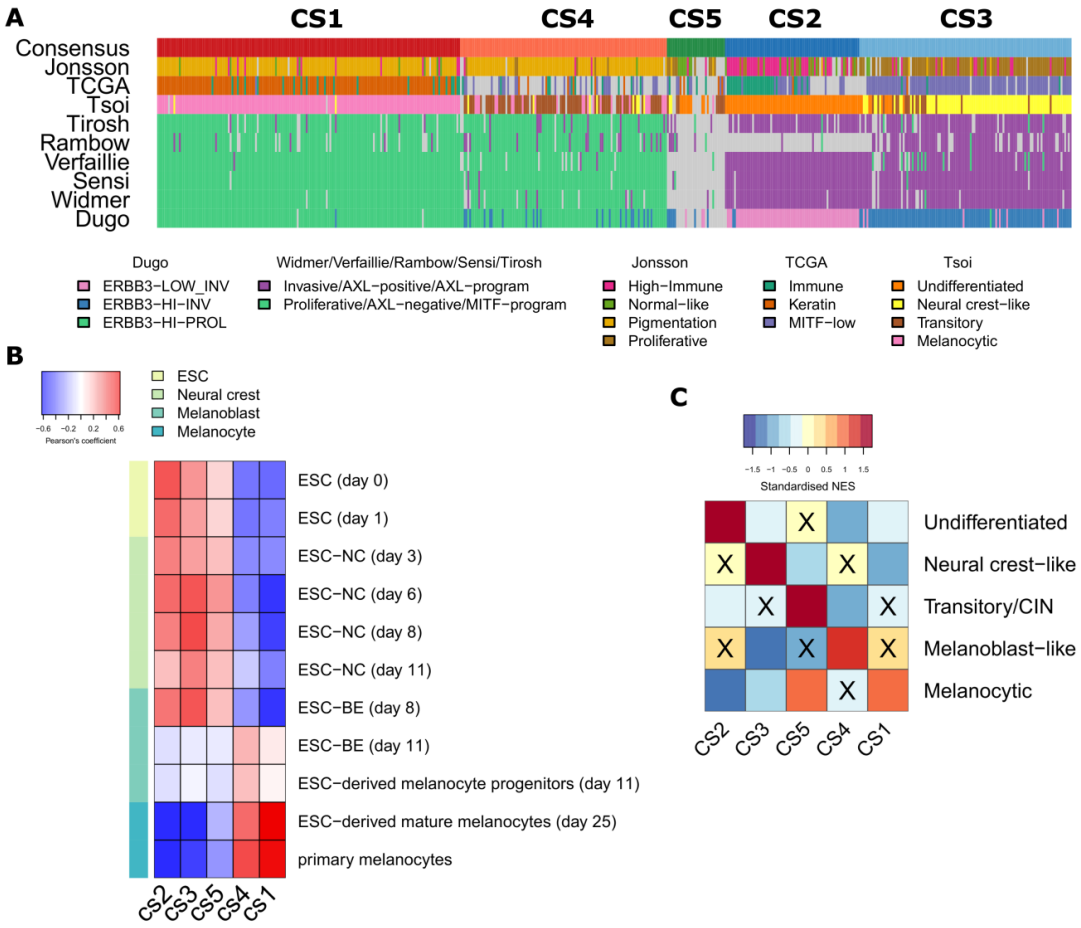

Figure 4.15. Similarity between *in vitro* and *in vivo* consensus subtypes. A) Heatmap showing the overlap between consensus subtypes and subtype labels assigned by the 9 melanoma gene signatures. B) Heatmap of the Pearson's correlation coefficient between the centroids of consensus subtypes and of experimental conditions of the melanocyte differentiation in vitro model from GSE45227. C) Enrichment of consensus subtype-specific signature derived from clinical tumours in the cell lines consensus subtypes.

4.7 Consensus subtypes and *in vitro* drug sensitivity

To link our consensus subtypes to drug sensitivity we analyzed gene expression data for the subset of 57 melanoma cell lines included in CCLE and matched pharmacological profiles for 481 drugs available from CTRP pharmacogenomic database. We and others previously showed that the invasive and proliferative phenotypes determined sensitivity to BRAF inhibitors with the first being intrinsically resistant to the drug and the latter being more sensitive (Konieczkowski et al., 2014; Muller et al., 2014; Dugo et al., 2015). We tested the response to BRAF inhibitors in the context of our new consensus

classification system, considering only 38 cell lines harboring the BRAF^{V600E} mutation. The same pattern of sensitivity across subtypes was observed for all four BRAF inhibitors tested (**Figure 4.16**). Undifferentiated cell lines were the most intrinsically resistant compared to all other groups. Interestingly, between the two proliferative subtypes the melanoblast-like was more sensitive compared to the melanocytic. The only cell line tested for the transitory/CIN subtype showed high sensitivity but unfortunately, we could not draw robust conclusions due to limited sample size.

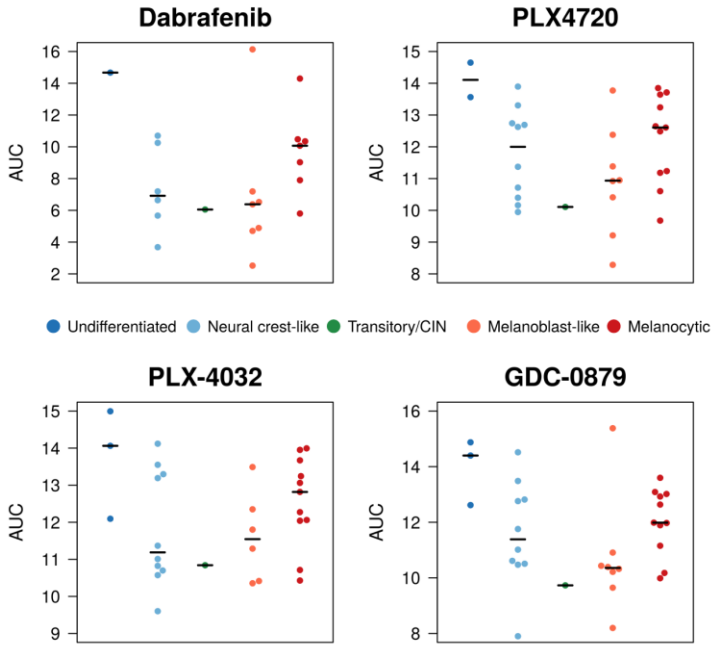

Figure 4.16. Subtype-specific sensitivity to BRAF inhibitors. Association between area under the curve (AUC) values for the four BRAF inhibitors available in CTRP database and consensus subtypes for CCLE cell lines. Higher AUC values are associated lower sensitivity. The horizontal black bars represent the median AUC value for each subtype. Not all cell lines were tested for all drugs.

Beyond BRAF inhibitors, we assessed the sensitivity level of our consensus subtypes to all other drugs tested, for all 57 cell lines irrespectively of BRAF mutations. This analysis was aimed to identify alternative drugs for BRAF inhibitor-refractory cells but also subtype-specific effective treatments for cell lines lacking the BRAF^{V600E} mutation. We compared the drug profiles of consensus subtypes in a pair-wise fashion and we selected 49 drugs with subtype-specific sensitivity (Tukey's post hoc p-value < 0.05 in at least one contrast). The pattern of sensitivity for these drugs was represented through a heatmap to identify the most effective drugs for each subtype (**Figure 4.17A**). The undifferentiated subtype was characterized by higher sensitivity to ferroptosis-inducing drugs such as ML210, ML162, erastin and 1S,3R-RSL-3, as already described (Tsoi et al., 2018). In addition, undifferentiated cell lines were more sensitive to dasatinib (SRC inhibitor), axitinib and Ki8751 (inhibitors of VEGFRs, c-KIT, and PDGFRs), SGX-523 (MET inhibitor),

NVP-TAE684 (ALK inhibitor) and CHIR-99021 (GSK3B inhibitor). The ROS-inducer BRD-K71935468, marinopyrrole A (putative inhibitor of MCL1), pifithrin-alpha (inhibitor of p53-dependent signaling) and RO4929097 (inhibitor of gamma-secretase) were more effective in neural crest-like cell lines. The melanocytic subtype showed higher sensitivity to several drugs belonging to the family of the histone deacetylase (HDAC) inhibitors and to the p38-MAPK inhibitor skepinone-L. On the other side, melanoblast-like cell lines showed higher sensitivity to several drugs including inhibitors of MEK1 and MEK2 (PD318088, selumetinib and trametinib) and proteasome inhibitors (MG-132 and bortezomib). The novel transitory/CIN subtypes showed specific sensitivity to the BIRC5 inhibitor YM-155 and to the inhibitor of PLK1 BI-2536 (**Figure 4.17B**) whose targets are up-regulated in this subtype (**Figure 4.17C**).

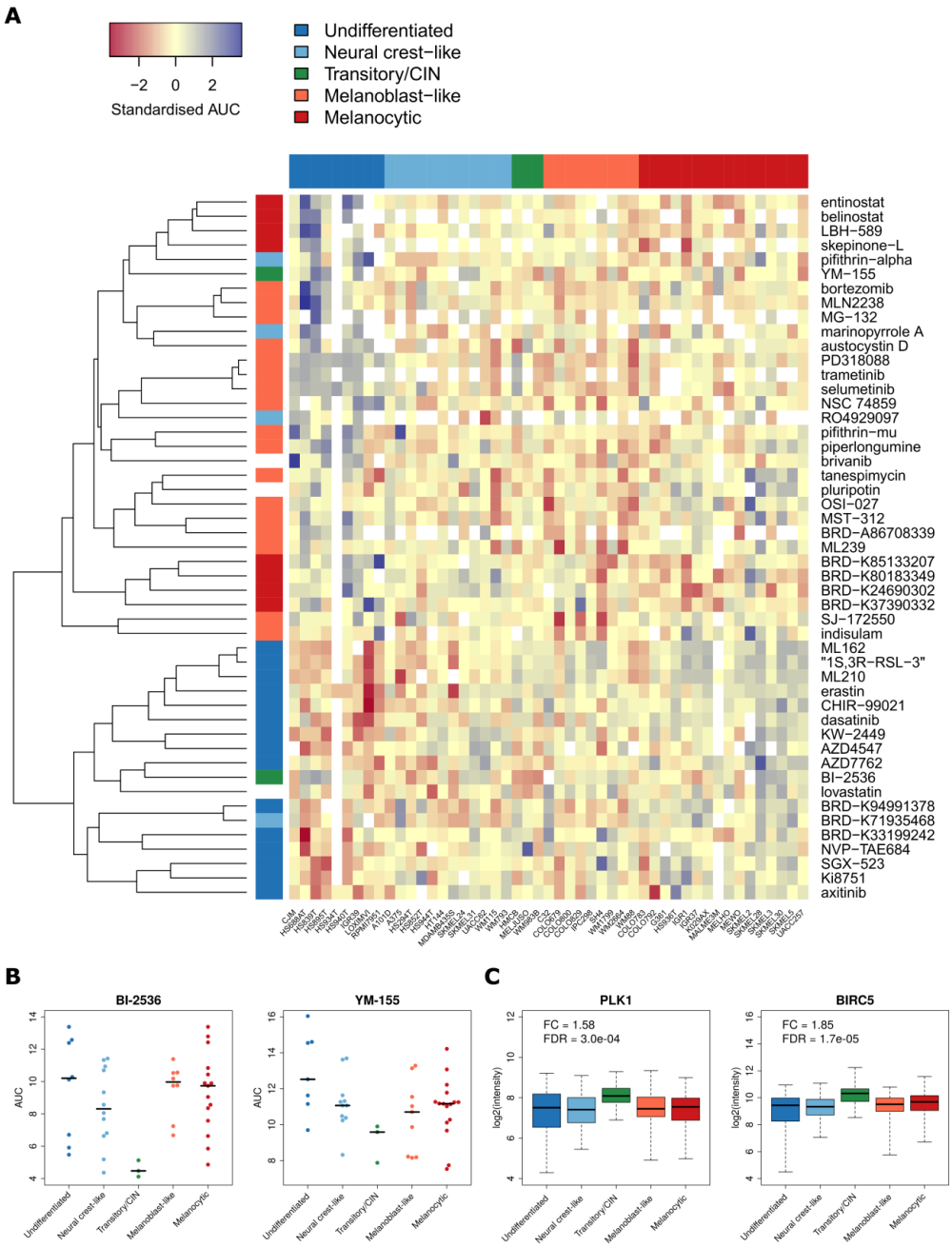

Figure 4.17. Association between drug response profiles and consensus subtypes. A) Heatmap of AUC values for CTRP drugs significantly associated to consensus subtypes. The colored bar on top delineates the consensus subtype. The vertical colored bar on the left indicates in which consensus subtype a drug is more effective. B) Plot showing the AUC values for the two drugs more effective in the transitory/CIN subtype. C) Box plot of the expression levels of the targets of the drugs represented in B). Box plot were obtained considering all 504 cell lines of the melanoma cell line dataset. Fold change (FC) and FDR refer to the contrast between transitory/CIN subtype and the others.

4.8 Predictive role of consensus subtypes in clinical tumours

To evaluate to which extent the role of consensus subtypes as predictive biomarker could be translated from an *in vitro* to an *in vivo* setting, we evaluated three gene expression datasets of BRAF-mutated metastatic melanoma patients treated with MAPK inhibitors (Rizos et al., 2014; Hugo et al., 2015; Kwong et al., 2015). Additionally, we evaluated the ability of consensus subtypes to predict response to immunotherapy in a

dataset of metastatic melanoma patients treated with the PD-1 inhibitor pembrolizumab (Hugo et al., 2016). For these analyses we focused on the prediction of intrinsic resistance, thus only pre-treatment samples were considered (**Table 4.3**).

Table 4.3. Gene expression datasets of metastatic melanoma patients treated with targeted- or immuno-therapy. Summary of datasets composition. N: number of pre-treatment patients; Dab: dabrafenib; Vem: vemurafenib; Tra: trametinib.

Dataset (N)	Sex	Age (Years)	Treatment	BRAF Mutation	Metastatic site
GSE50509 (21)	M (14) F (7)	54 ± 18	Dab (16) Vem (5)	V600E (14) V600K (6) V600R (1)	Subcutaneous (14) Lymph node (3) Brain (4)
GSE65185 (18)	M (13) F (5)	58 ± 13	Dab (1) Vem (12) Dab+Tra (4) Vem+GDC0973 (1)	V600E (13) V600K (4) V600R (1)	Cutaneous (7) Subcutaneous (9) Lymph node (1) Breast (1)
EGAD00001001306 (14)	NA	51 ± 17	Vem (3) Dab + Tra (11)	V600E (14)	NA*
GSE78220 (26)	M (18) F (8)	61 ± 13	Pembrolizumab§ (26)	V600E (8) V600K (2) V600R (1) WT (13)	Subcutaneous (14) Lymph node (2) Other (10)

For patients treated with anti-MAPK therapy, we applied GSEA to correlate the expression of the consensus subtype-specific signatures to the best overall response evaluated according to RECIST criteria (Eisenhauer et al., 2009). For the anti-PD-1 dataset, consensus subtypes signatures were evaluated in association to the classification of patients in complete response (CR), partial response (PR) or progressive disease (PD). We found an inconsistent pattern of enrichment among datasets. The undifferentiated signature was positively enriched in patients with the strongest disease reduction in GSE50509 and showed a trend of positive enrichment in the other two dataset of targeted therapy (**Figure 4.18A**). This is in contrast to what was observed *in vitro*, where undifferentiated cell lines where associated to drug tolerance. The melanoblast-like signature was significantly negatively correlated to response to MAPK inhibitors in GSE50509 and EGAD00001001306, while showed an opposite behavior in GSE65185.

For the immunotherapy dataset we observed a positive enrichment of the melanoblast-like, and a negative enrichment of the melanocytic and the transitory/CIN

signatures in CR patients compared to PR or PD patients. Only the melanoblast-like signature maintained its significantly positive enrichment in PR patients compared to PD (**Figure 4.18A**). Additionally, we assessed the association to therapeutic response for a very recently published gene sets derived from single-cell RNA-sequencing of melanoma tumours during treatment with RAF inhibitors (Rambow et al., 2018). As observed for our consensus signatures, the pattern of enrichment is not reproducible across all datasets (**Figure 4.18B**).

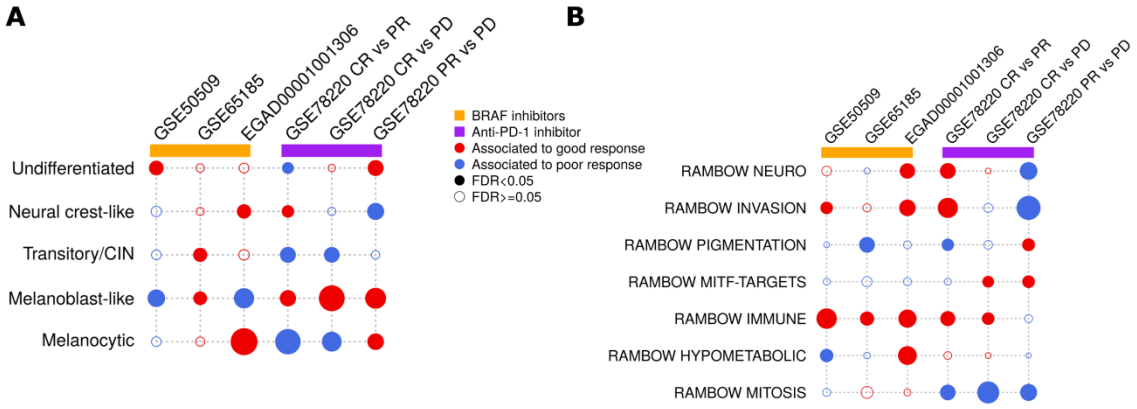

Figure 4.18. Assessment of the predictive value of melanoma transcriptional state signatures in clinical tumours. A-B) Bubble plots showing the association of selected signatures and response to anti-MAPK and anti-PD-1 therapy in publicly available datasets. The size of each circle is proportional to the normalized enrichment score calculated by GSEA.

To get a broader view of the pathways associated to therapeutic response we performed GSEA using the HALLMARK collection, a list of gene sets representing well-defined biological processes (**Figure 4.19**). Again we observed a certain level of inconsistency across datasets but gene sets related to interferon-alpha and interferon-gamma signaling were found to be positively associated to response for all three datasets of targeted therapy and for CR patients compared to PR or PD in the anti-PD-1 dataset. Other immune-related gene sets such as inflammatory response and IL6-JAK-STAT3 signaling showed a trend of positive association in all datasets under investigation.

These results suggest that the consensus subtypes alone could not predict *in vivo* the response to targeted MAPK inhibition and PD-1 blockade but other factors such as the cross-talk between tumour and immune and stromal cells or the presence of additional genomic alterations could impact on the therapeutic response. However, we could not exclude that the poor performance of consensus subtypes in predicting response and the poor reproducibility across datasets arose from the small sample size of available datasets and their heterogeneity in terms of sample composition, therapeutic treatment and

technological platform used for the transcriptional profiling. Analysis of much larger and homogenous cohorts is required for a robust assessment of molecular subtypes of melanoma as predictive biomarkers.

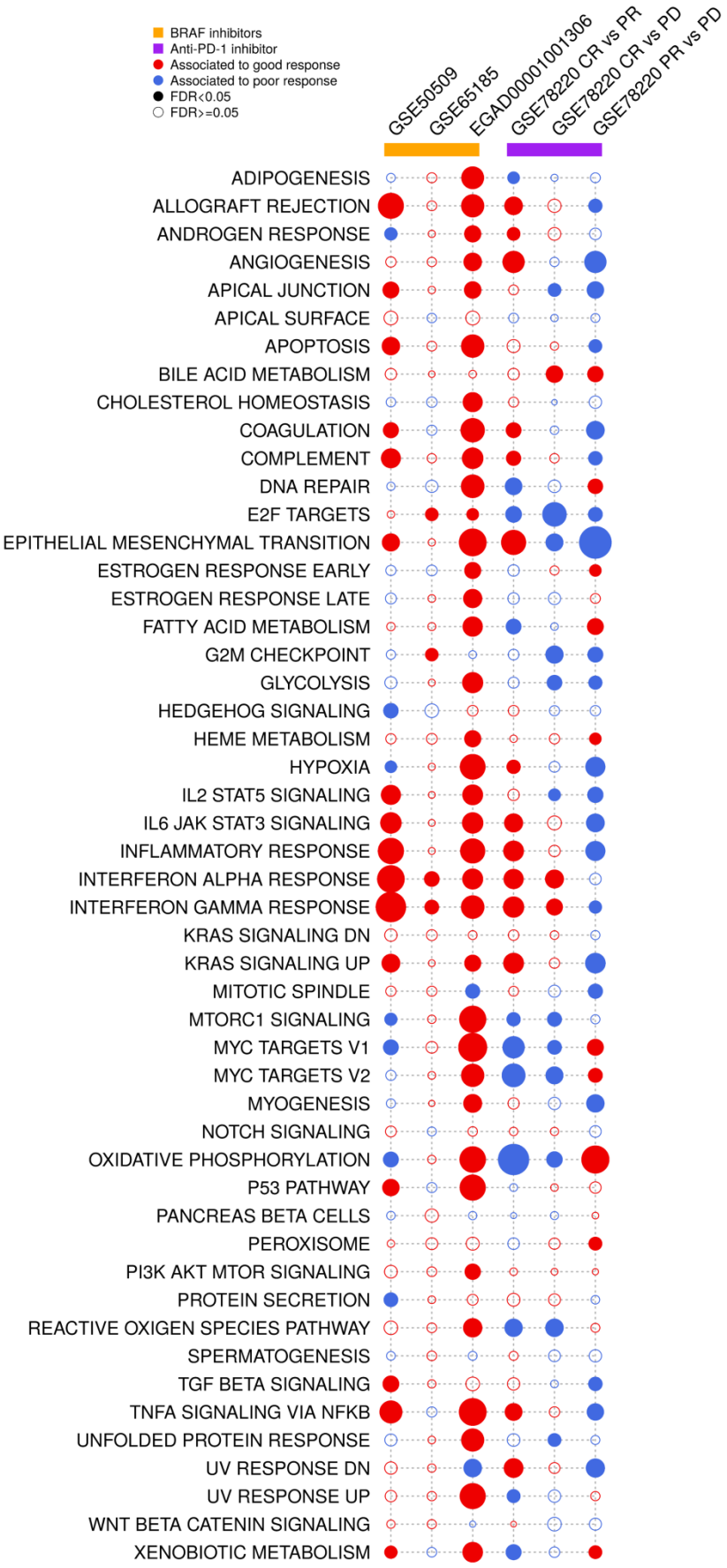

Figure 4.19. Biological processes associated to targeted- or immunotherapy response. The size of each circle is proportional to the normalized enrichment score calculated by GSEA.

4.9 Multi-omics characterization of consensus subtypes

4.9.1 Somatic mutations and copy number aberrations

According to the mutational pattern of the most frequently mutated genes in melanoma (BRAF, RAS and NF1) melanoma can be classified in four genomic subtypes: mutant BRAF, mutant RAS, mutant NF1 or triple WT if none of the 3 genes is mutated. Here we assessed whether an association between consensus subtypes and genomic subtypes existed. We found that beside a higher frequency of BRAF-mutant tumours in the neural crest-like group, the four genomic subtypes were distributed across the five consensus subtypes indicating that the driver mutated genes in melanoma did not contribute to the determination of molecular subtypes at the transcriptional level (**Figure 4.20A**). We next investigated the overall mutational pattern of melanoma samples. The five consensus subtypes showed a comparable mutational burden (**Figure 4.20B**). The combination of somatic mutations in a tumour can be deciphered in several mutational signatures arising from different mutational processes and previously identified through genomic analysis of multiple cancer types (Alexandrov et al., 2013b; Alexandrov et al., 2013a). Many of these signatures are cancer-specific and melanoma is characterized by a mutational signature enriched of CC>TT mutations at dipyrimidines photodimers, typical of UV damage. We measured the contribution of 30 mutational signatures from COSMIC in TCGA-SKCM dataset and we assessed whether any signature was enriched in a particular consensus subtype. Three major signatures were identified to contribute to the mutational pattern of TCGA melanoma samples. As expected, the most abundant signature was the UV-related Signature 7, followed by Signature 1 that is recurrent in all cancer types and is correlated to age. Finally, we found Signature 11, another signature enriched in melanoma with a pattern of mutation similar to that of alkylating agents. None of the signatures was enriched in a specific consensus subtype (**Figure 4.20C**). We finally tested whether other genes beside BRAF, RAS and NF1 were preferentially mutated in any of the consensus subtype. After multiple testing correction no statistically significant associations were observed. Inspection of the top-associated mutated genes (Fisher's exact test nominal p-value < 0.001) highlighted that their trend of association with consensus subtypes could be due chance and not biologically relevant as these mutations occurred in a small fraction of patients (**Figure 4.20D**).

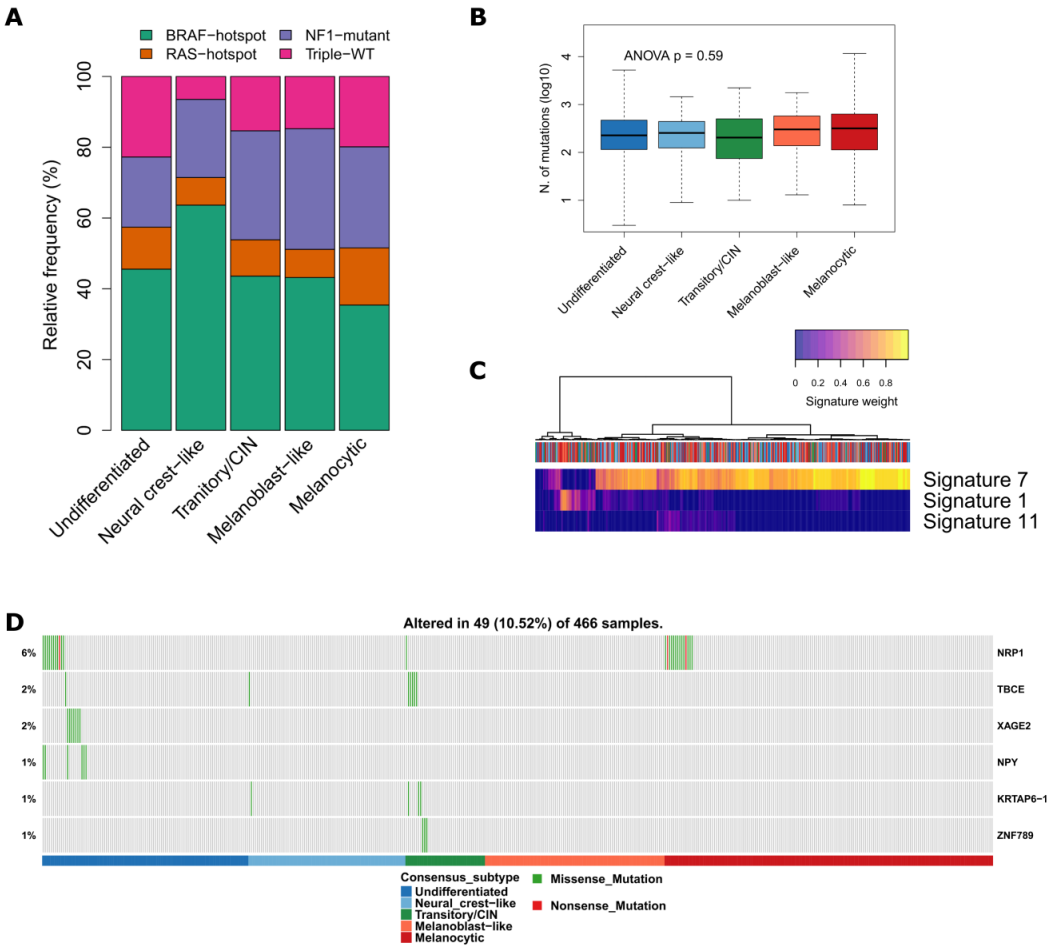

Figure 4.20. Somatic mutation analysis of TCGA-SKCM consensus subtypes. A) Distribution of genomic subtypes across the five consensus subtypes. B) Mutational load of each subtype measured as the number of somatic mutations for each sample on a log10 scale. C) Heatmap showing the relative contribution of signatures 1, 7 and 11. The color bar above the heatmap represents the consensus subtype of each sample. D) Oncoplot showing the frequency of mutations across subtypes for the genes showing a trend of association with consensus subtypes ($p < 0.001$). Only samples with a mutation in at least one of the genes are shown.

For the same patients we analyzed copy number aberration data. We first assessed whether the consensus subtypes differed in terms of tumour cell ploidy and purity, inferred from copy number data. Consensus subtypes showed comparable level of tumour ploidy (**Figure 4.21A**) but showed different tumour purity with the undifferentiated subtype characterized by the lowest purity. This result was in agreement with the tumour purity estimated from gene expression data with the ESTIMATE algorithm (**Figure 4.11C**) and confirmed the high level of stromal component within these tumours.

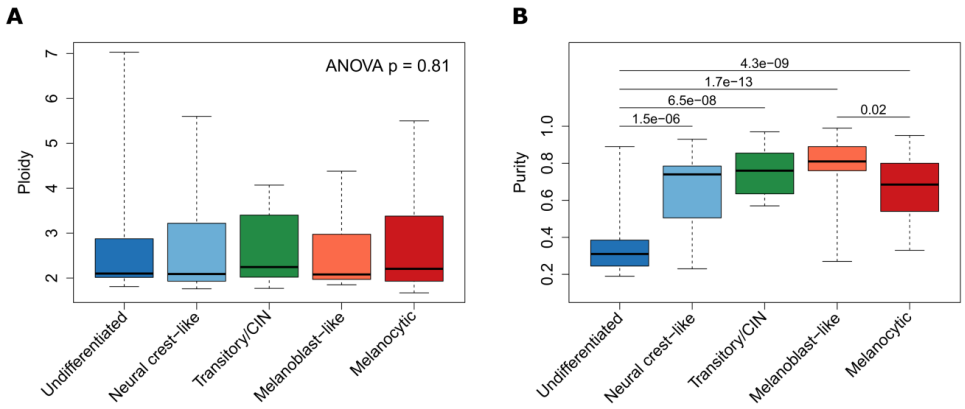

Figure 4.21. Tumour heterogeneity based on copy number data. A-B) Box plots showing the ploidy (A) and tumour purity (B) estimated by ABSOLUTE algorithm.

By counting the number of amplifications and deletions we obtained an estimate of the chromosomal instability of each tumour (**Figure 4.22A**). We found that the transitory/CIN subtype had on average a higher number of amplifications and deletions, supporting the genomically instable phenotype of this subtype. In contrast, the undifferentiated subtype was characterized by a lower number of aberrations and especially of deletions. However, given the high content of non cancerous cells we cannot exclude that in this group several copy number aberrations could be missed due to a loss of sensitivity arising from the presence of normal stromal cells. Similarly to mutations, we aimed to identify copy number events associated to consensus subtypes. Using GISTIC algorithm we first identified significantly recurrent amplifications and deletions across the entire TCGA-SKCM cohort. Several genomic regions were found to be significantly altered, divided in 17 amplified and 24 deleted regions (**Figure 4.22B**). Regardless of consensus subtype, we found recurrent deletions of CDKN2A (9p21.3) and PTEN (10q23.31) and amplification of TERT (5p15.33), MITF (3p13), MYC (8q24.3) and BRAF (7q34). Focusing on the significantly altered regions we tested their association with consensus subtypes. Out of the 41 regions tested we found statistically significant association for 21 of them. For example, we found that deletions at 19p13.3 and 11q23.3 occurred preferentially in the melanocytic subtype. Deletions of 10q23.31 and 9p21.3 regions, containing PTEN and CDKN2A, respectively, were more recurrent in the undifferentiated subtype. However, similarly to somatic mutations, despite these regions were statistically significantly more frequent in specific subtypes, we observed that they were recurrent across all of them (**Figure 4.22C**).

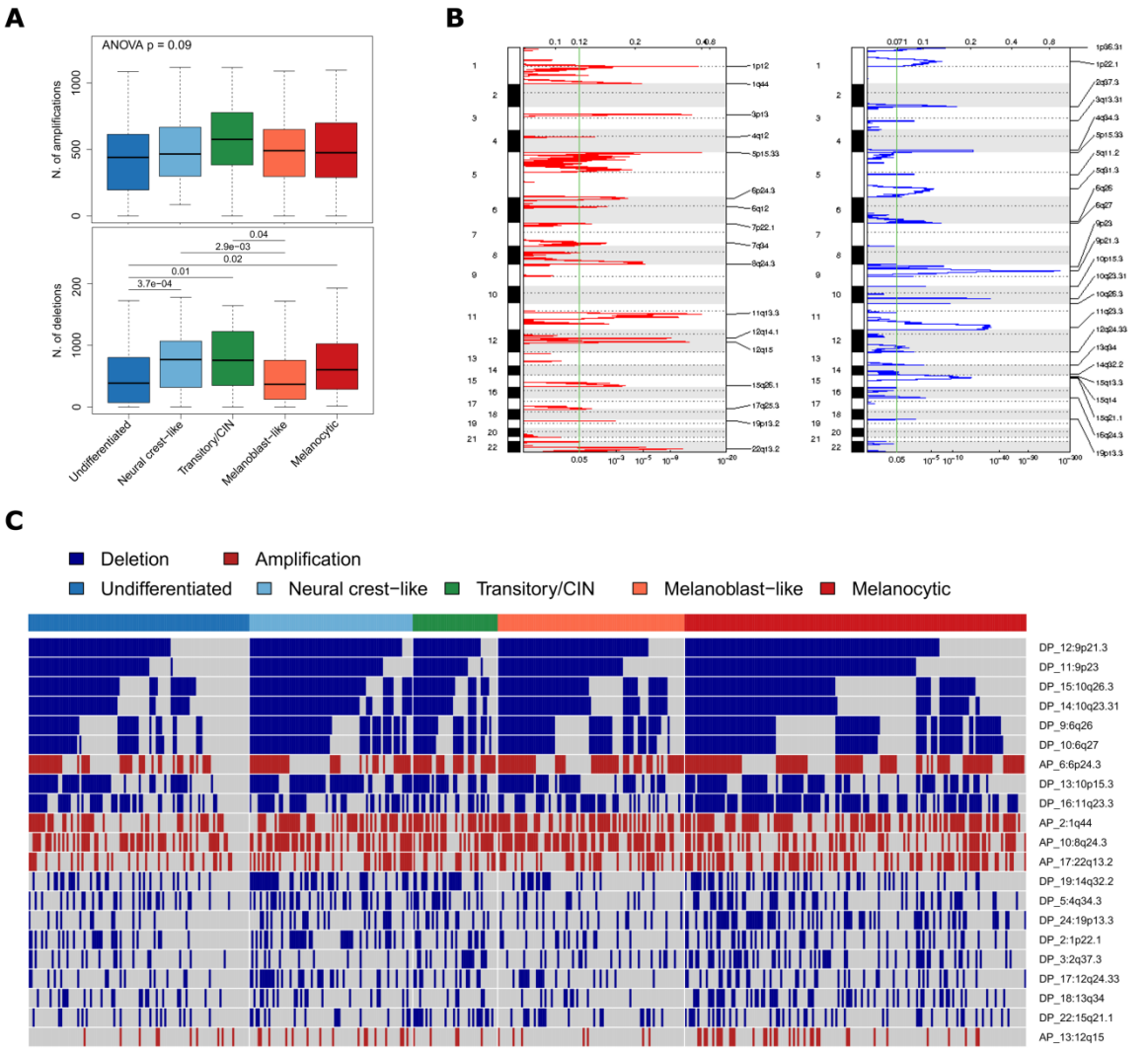

Figure 4.22. Somatic copy number aberration analysis of TCGA-SKCM consensus subtypes. A) Number of amplified and deleted genes identified by GISTIC for samples belonging to different consensus subtypes. B) GISTIC plots showing the most recurrently amplified (red) and deleted (blue) regions in the entire TCGA-SKCM cohort, regardless of consensus subtype. The right bar shows the chromosomal location and on the left are reported the regions significantly altered (FDR < 0.05, vertical green line). C) Oncoplot reporting the frequency of amplifications (red) and deletions (blue) of the regions found significantly associated with consensus subtypes (FDR < 0.05).

Overall, these results showed that somatic mutations and copy number aberrations did not play a critical role in the determination of the consensus subtypes. To confirm this hypothesis we performed an integrative analysis of gene expression and copy number data. We first correlated the expression of each gene with its copy number measurement and the distribution of the Pearson's correlation coefficients showed a shift of the distribution toward positive values (**Figure 4.23A**). This suggested that the expression of several genes was influenced by their copy number status. However, according to GSEA, we did not observe any significant enrichment toward the positively correlated genes for the consensus subtypes signatures and for GO biological process

terms representative of the main pathways altered in consensus subtypes (**Figure 4.23B-C**).

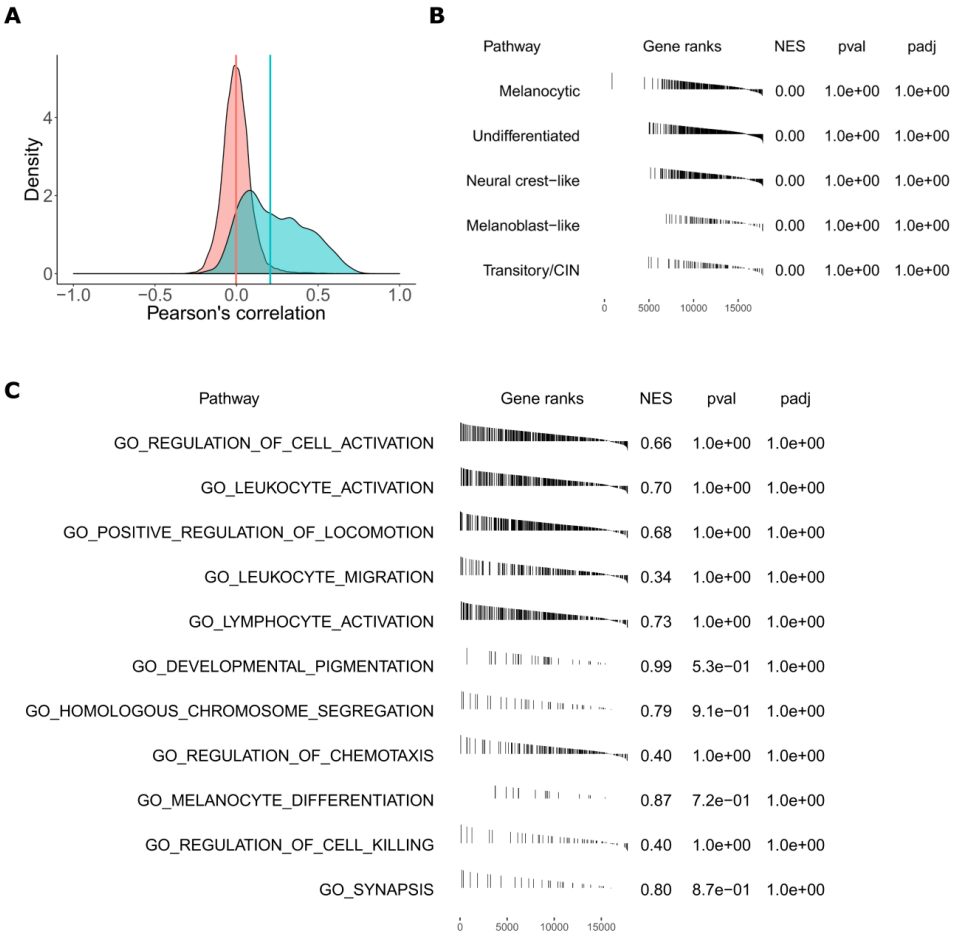

Figure 4.23. Integrative analysis of RNA-Seq and copy number data. A) Distribution of the Pearson's correlation values the expression of each gene and its copy number log2ratio. The green curve is the observed distribution, compared to a random distribution obtained by randomly shuffling the genes 1000 times. B-C) Enrichment plots for selected gene sets tested against the list of genes pre-ranked according to the Pearson's correlation coefficient. The "Gene ranks" column shows the location of the genes included in each pathway along the pre-ranked gene list. NES: normalized enrichment score; padj: adjusted p-value.

4.9.2 DNA methylation

Gene transcription is heavily influenced by epigenetic modification such as DNA methylation. To describe the epigenetic landscape of consensus subtypes we analyzed methylation data for TCGA-SKCM samples. Methylation profiles were obtained using Illumina Infinium HumanMethylation450K BeadChip arrays that target more than 450,000 CpG methylation sites along the genome.

To get a first overview of the methylation status of samples belonging to the different consensus subtypes we first performed an explorative analysis by simply visualizing the mean methylation level of each sample in function of consensus subtypes. The only significant difference was observed between the undifferentiated and neural crest-like subtypes with the latter showing a decreased level of methylation (**Figure**

4.24A). We next identified differentially methylated CpG sites comparing each subtype with all other four. More than 68000 CpG sites were significantly differentially methylated across subtypes (FDR < 0.05 and β fold change > 0.1) mapping on 11673 unique genes. However, differentially methylated CpGs were not equally distributed across subtypes but were prevalently found in the undifferentiated and melanoblast-like subtypes (**Figure 4.24B**). Among the most significantly hyper-methylated sites in the undifferentiated subtype we found CpGs located in the nearby region of transcription start site of MITF-M transcript variant (**Figure 4.24C**), which is specifically expressed in melanocytes.

Due to the high number of differentially methylated sites and the one-to-many relationship between genes and CpG sites, a full interpretation of the results is made extremely complex. Indeed, many genes have multiple CpG methylation sites distributed in nearby regions, inside the promoter and in the gene body and the effect on transcription is unknown for the majority of them. Thus, we applied a more general approach to understand whether the transcriptional changes of consensus subtypes are under epigenetic control. For this purpose, we calculated a single methylation value for each gene by averaging the intensities of the probes mapping in the promoter regions and we correlated these values to gene expression levels. Overall, we found that the observed distribution of the Pearson's correlation coefficients for matched methylation and transcriptomics data was shifted toward negative values (median = -0.08) (**Figure 4.25A**). A significant negative correlation between gene expression and methylation level was observed for 6838 genes. However, 1847 genes showed a significantly positive correlation. According to GSEA, genes found differentially expressed across consensus subtypes were enriched toward the genes negatively correlated with their methylation status (**Figure 4.25B**). Moreover, several biological processes previously described as key altered pathways in the consensus subtypes were enriched of anti-correlated genes (**Figure 4.25C**). These findings suggested that the expression of genes relevant for the biology of consensus subtypes was under DNA methylation control.

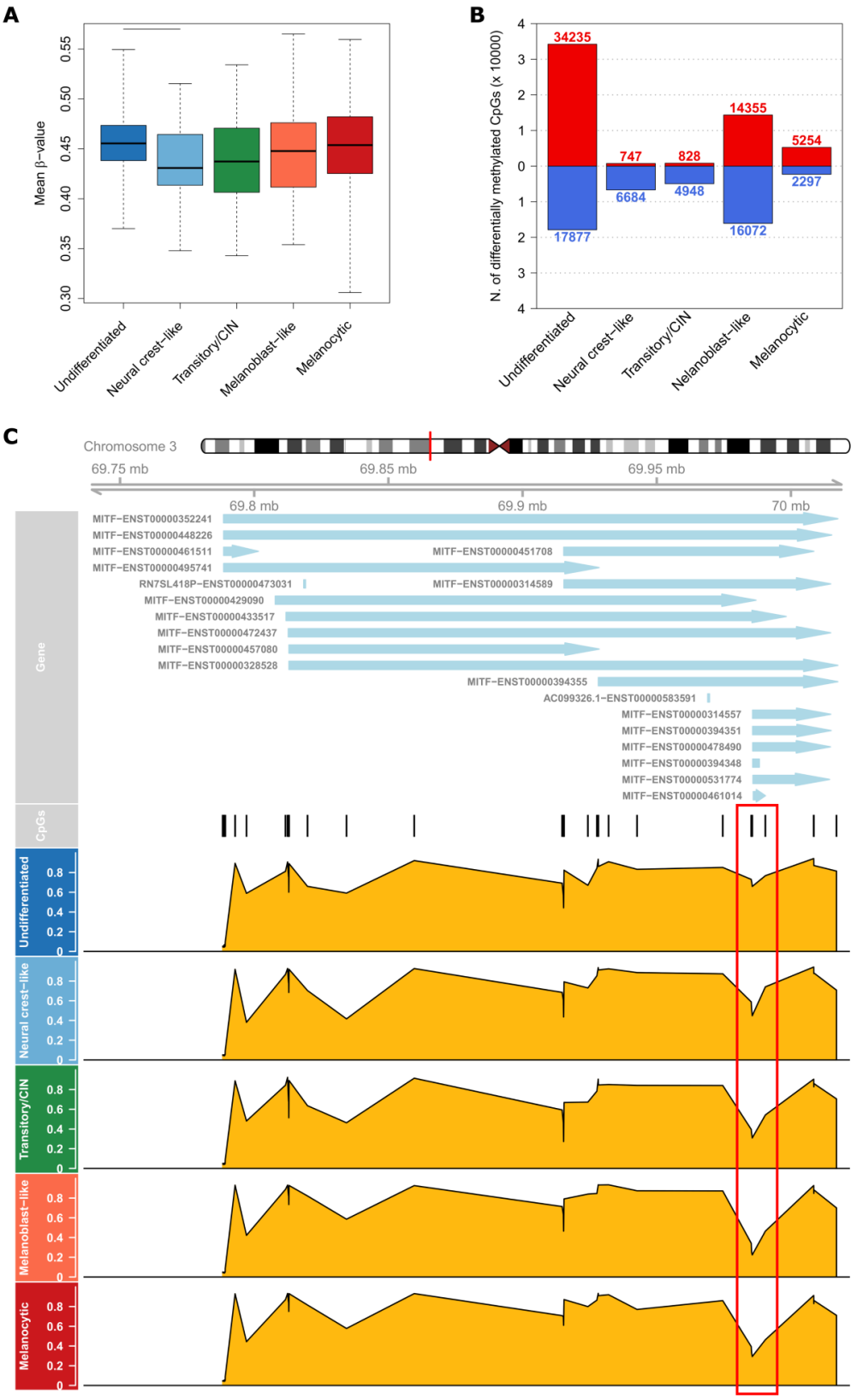

Figure 4.24. DNA methylation analysis of TCGA-SKCM consensus subtypes. A) Average methylation levels of samples within each consensus subtypes. B) Number of significantly differentially methylated CpG sites (FDR < 0.05 and absolute β-value fold change > 0.1). Red bars denote methylated CpGs, blue bars demethylated CpGs. C) MITF methylation status. The top panel shows MITF locus on chromosome 3 (vertical red bar) and all MITF transcript isoforms from Ensembl. The black vertical bars in the middle panel show the location of CpG sites interrogated by Illumina Infinium450 arrays. The lower panel shows the average β-value of the CpG sites along MITF gene in each subtype. The red box highlights the differentially methylated CpG probes targeting the nearby region of the transcription start site of MITF-M isoform (MITF-ENST00000394351).

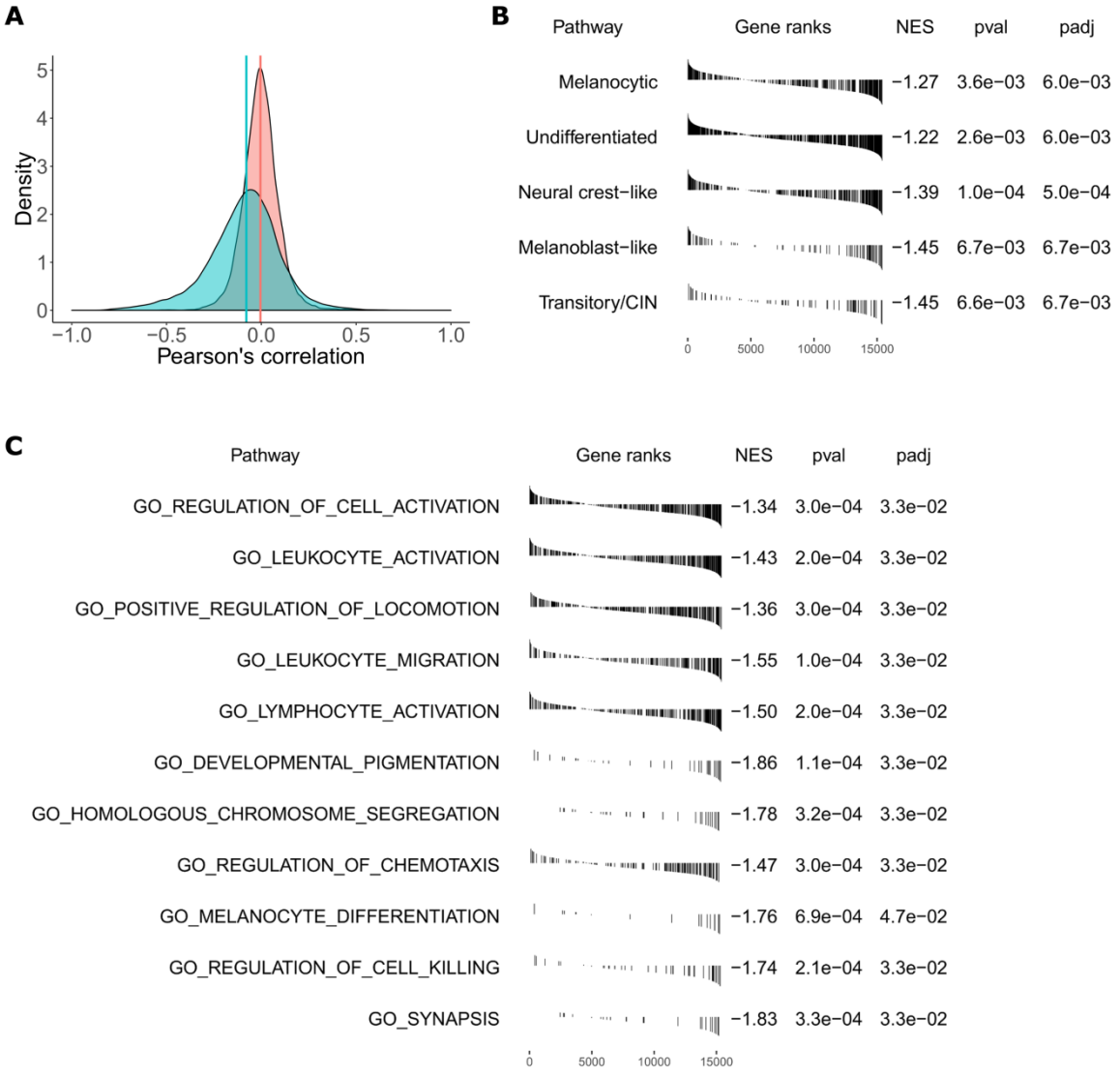

Figure 4.25. Integrative analysis of RNA-Seq and methylation data. A) Distribution of the Pearson's correlation values the expression of each gene and its methylation level. The green curve is the observed distribution, compared to a random distribution obtained by randomly shuffling the genes 1000 times. B-C) Enrichment plots for selected gene sets tested against the list of genes pre-ranked according to the Pearson's correlation coefficient. The "Gene ranks" column shows the location of the genes included in each pathway along the pre-ranked gene list. NES: normalized enrichment score; padj: adjusted p-value.

4.9.3 microRNA expression profiling

MicroRNAs (miRNAs) are small non-coding RNAs implicated in post-transcriptional regulation of gene expression through translation inhibition or mRNA target degradation. They participate in the regulation of many cellular processes and they are key players in many human diseases, including cancer. In addition to cell-intrinsic activity they also mediate inter-cellular communication between tumour cells and their microenvironment. Given the high level of transcriptional regulation observed across the consensus subtypes we analyzed miRNA expression data of TCGA-SKCM samples to define sets of miRNAs involved in subtype determination through gene expression regulation. Comparing each

subtype with all other four we found many miRNAs significantly differentially expressed (FDR < 0.05, **Figure 4.26**).

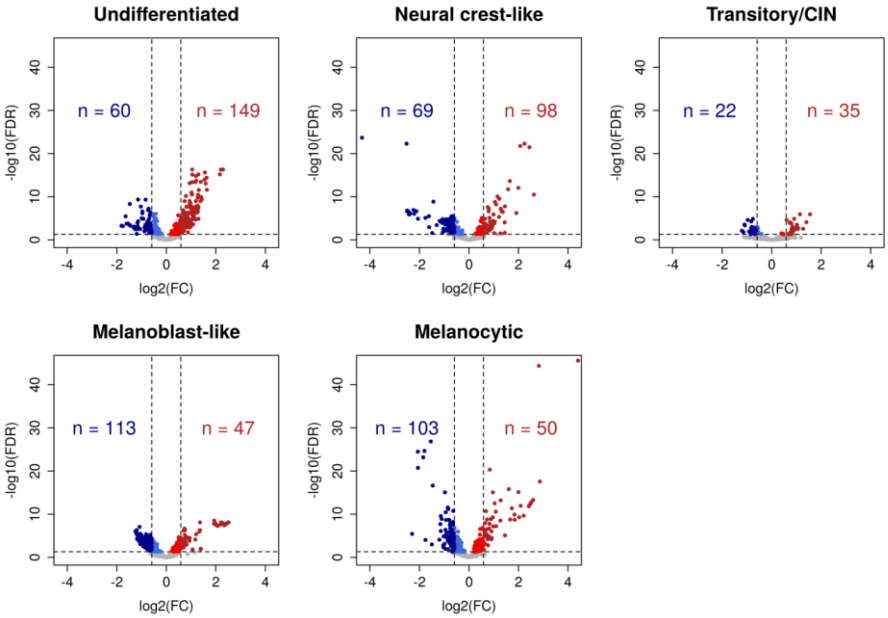

Figure 4.26. miRNA expression analysis of TCGA-SKCM consensus subtypes. Volcano plots showing the number of differentially expressed miRNAs in each subtype. The x-axis represents the log2 fold change, the y-axis the –log10 of the false discovery rate. miRNAs with an FDR < 0.05 (horizontal dashed line) are highlighted. Significant miRNAs with an absolute fold change ≥ 1.5 (vertical dashed lines) are highlighted with a darker color. Red: up-regulated miRNAs; blue: down-regulated miRNAs.

Given the high number of differentially expressed miRNAs and the extreme complexity of the miRNA-mRNA regulatory dynamics (a miRNA can target multiple mRNAs and an mRNA can be targeted by several miRNAs) we applied an integrative approach to define the biological role of differentially expressed miRNAs in melanoma. We first selected miRNAs and genes significantly differentially expressed across subtypes with a fold change of at least 1.5 and 2 in at least one contrast, respectively. Then using *in silico* miRNA target predictions and matched gene-miRNA expression data we identified a list of candidate targets for each miRNA. Candidate targets for a miRNA were those predicted by at least 2 out of 6 prediction algorithms used and negatively correlated to their targeting miRNA (Pearson's correlation coefficient < -0.2). Each list of candidate targets was then tested for significantly enriched pathways (FDR < 0.05). Finally, we identified clusters of miRNAs involved in the same pathways using hierarchical clustering and Jaccard index as similarity measure. Using this approach we found seven clusters of miRNAs targeting genes involved in the same biological processes (**Figure 4.27A**). To assign a biological function to each cluster we investigated the pathways more frequently enriched in the target list of each miRNA (**Figure 4.27B**). Clusters 1, 2 and 6 included the majority of miRNAs down-regulated in the undifferentiated and neural crest-like

subtypes. The anti-correlated targets of these miRNAs were enriched in immune-related processes and extra-cellular matrix remodeling, in line with the observation from gene expression that the two invasive subtypes over-expressed immune and stromal genes. On the other side, clusters 4, 5 and 7 contained miRNAs up-regulated in the undifferentiated and neural-crest subtypes whose targets were implicated in pigmentation and melanocyte differentiation. Finally, cluster 3 was populated of miRNAs characterized by a poor degree of shared pathways both with miRNAs of the same cluster and with all other miRNAs.

For each cluster we focused our attention on the top-5 miRNAs most differentially expressed across consensus subtypes and for each miRNA we reported the top-5 anti-correlated targets (**Figure 4.28**). Among the 15 top differentially expressed miRNAs of clusters 1, 2 and 6, up-regulated in melanoblast-like and melanocytic subtype, 11 were members of the miR-506-514 cluster (hsa-miR-508-5p, hsa-miR-508-3p, hsa-miR-509-5p, hsa-miR-509-3p, hsa-miR-510-5p, hsa-miR-514a-3p and hsa-miR-514b-3p) that along with hsa-miR-211-5p were shown to be melanocyte-specific miRNAs (Boyle et al., 2011; Streicher et al., 2012). Among the top differentially expressed miRNAs of clusters 4, 5 and 7, up-regulated in the undifferentiated and neural crest-like subtypes we found members of the miR-143/145 cluster (hsa-miR-143-5p and hsa-miR-145-5p), miR-199/214 cluster (hsa−miR−199a−5p, hsa−miR−214−3p), miR-100 cluster (hsa−miR−125b−5p and hsa−let−7a−2−3p) and miR-181 cluster (hsa−miR−181d−5p and hsa−miR−181c−3p). MITF and several of its target genes such as BCL2, MBP, CAPN3 and TYR were among the top anti-correlated predicted targets of these miRNAs. Other members of miR-199/214 and miR-100 clusters were among the top-5 miRNAs up-regulated in the two invasive subtypes and fell in cluster 3. Among the top anti-correlated targets of hsa-199a-3p, hsa-miR-199b-3p and hsa-miR-214-5p there was TFAP2A. It was previously shown that miR-214 increases migration and invasion of melanoma cells by silencing TFAP2A and TFAP2C (Penna et al., 2013). In addition these miRNAs are anti-correlated with well known MITF targets such as MLANA, MLPH and VAT1. CABLES1, an important regulator of cell cycle, is targeted by many miRNAs up-regulated in the undifferentiated and neural crest-like subtype.

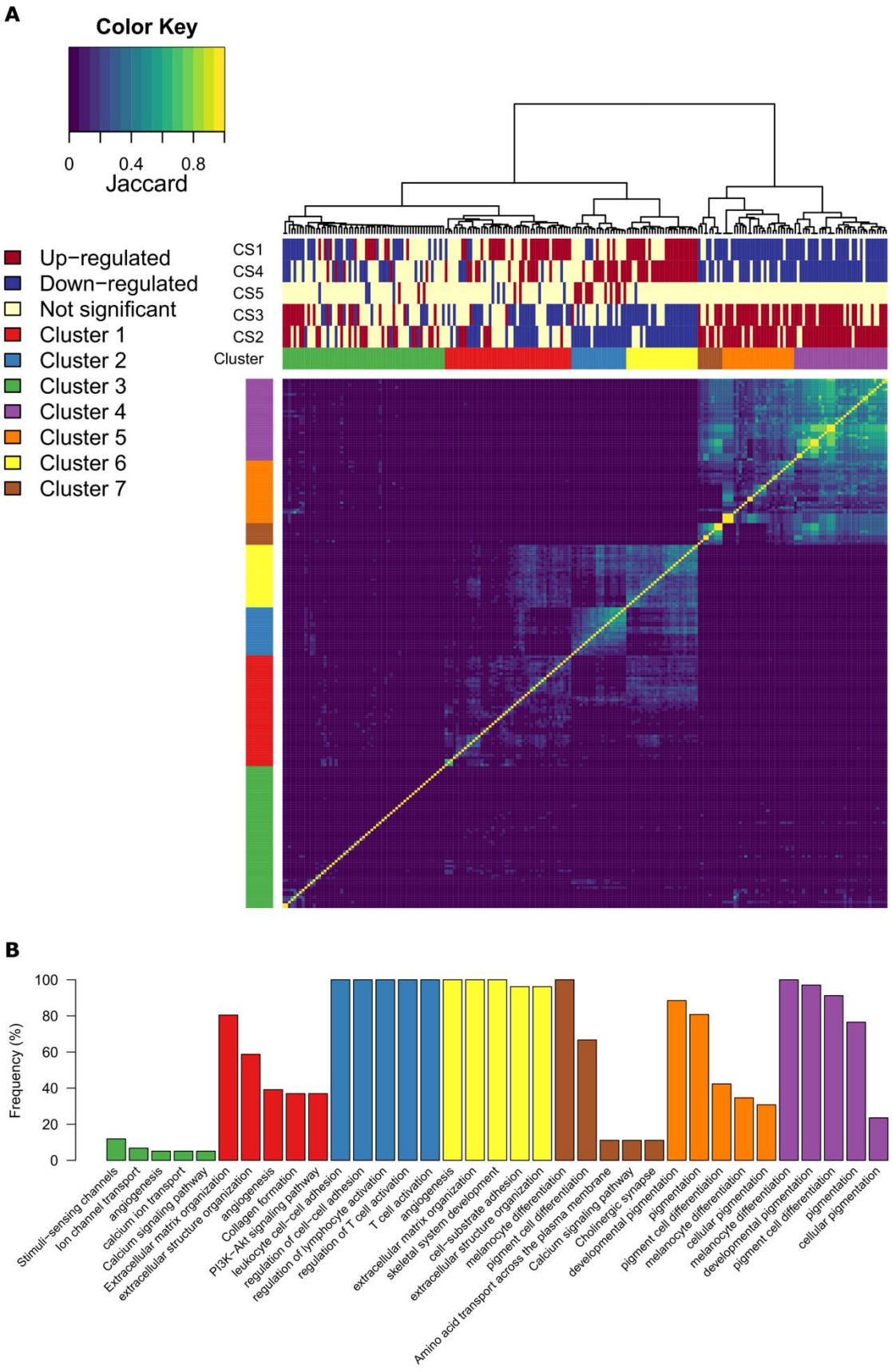

Figure 4.27.Pathway analysis of miRNAs differentially expressed across consensus subtypes. A) Heatmap of the Jaccard index values calculated for each pair of miRNAs on the basis of the shared pathways. Column and row dendrograms were obtained by hierarchical clustering of the Jaccard indices (Euclidean distance, Ward linkage). The color bars on the top represent the 7 identified clusters and the up- or down-regulation of the miRNAs in each consensus subtype. B) Bar plot of the top-5 most frequently enriched pathways (FDR < 0.05) in each cluster of miRNAs.

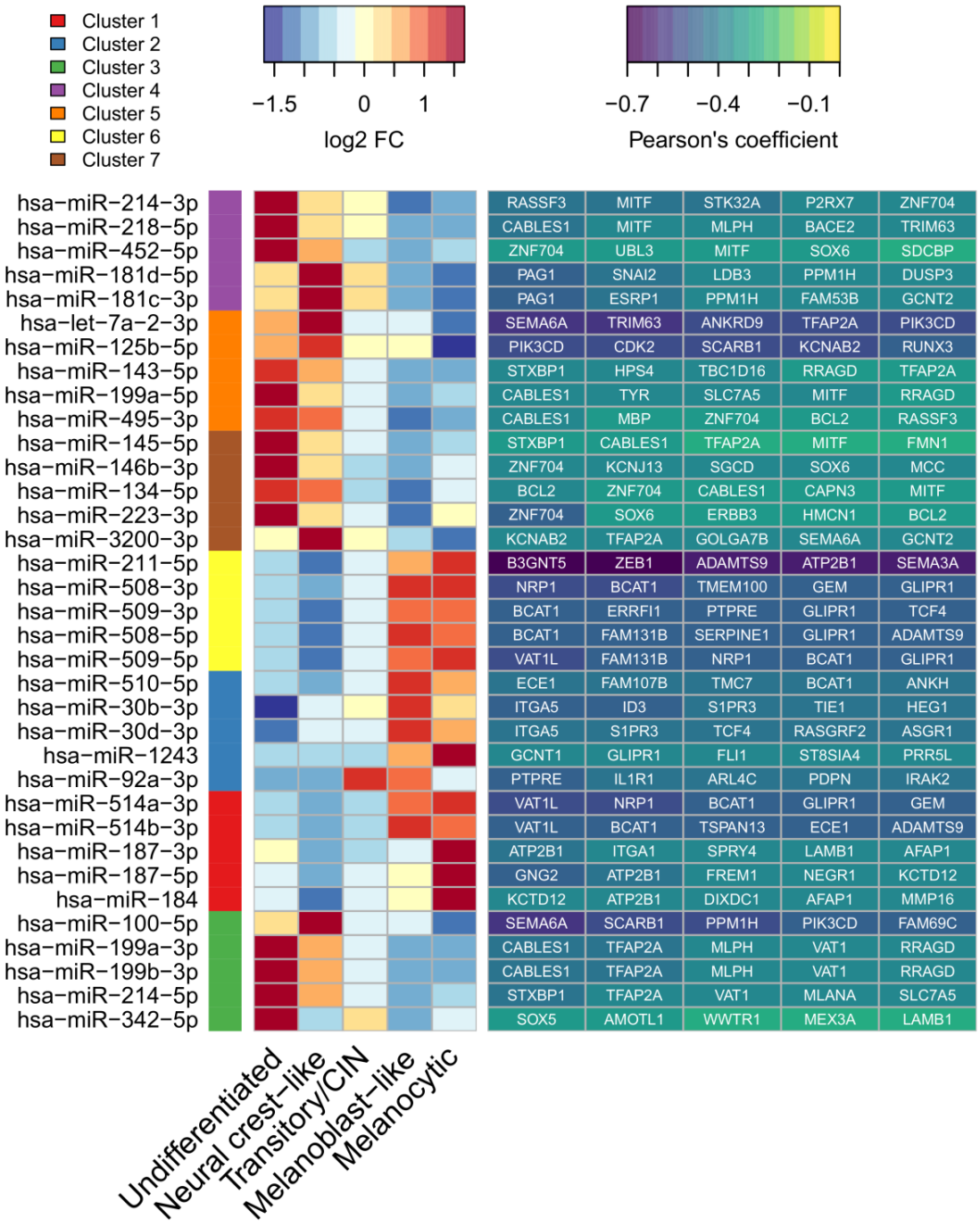

Figure 4.28. Top-differentially expressed miRNAs and related targets. The heatmap on the left shows the log2 fold changes across consensus subtypes of the top-5 most significantly differentially expressed miRNAs for each cluster identified in Figure 15A. The heatmap on the right report the top-5 anti-correlated predicted target genes of the miRNAs on the left. The color scale is proportional to the Pearson's correlation coefficient of the miRNA-target pair.

5 DISCUSSION

In this thesis we assessed the consistency of nine melanoma gene expression signatures through a harmonized classification of large datasets of melanoma clinical tumours and cell lines and we derived the consensus molecular subtypes of melanoma. Gene expression-based classifications in melanoma have been previously reported and led to the identification of up to four melanoma subtypes, distinguished by proliferation rate, invasive capability and differentiation status. In addition to technological and analytical differences, these signatures were derived from *in vitro* cell lines, clinical tumours or a combination of both, thus increasing the level of complexity. The overlap and robustness of these classifications remains unknown.

5.1 Identification and characterization of melanoma consensus subtypes

Lauss and colleagues (Lauss et al., 2016) already carried out an attempt to compare different melanoma classifications, even if they focused only on two subtyping schemes. They compared their melanoma subtypes (Jonsson et al., 2010) with those of TCGA (The Cancer Genome Atlas Network, 2015) and they found a certain degree of overlap in the classifications despite the two gene signatures were poorly overlapping. When we measured the overlap between the nine gene signatures considered in our work we found similar results. No genes were found to be represented in all nine signatures and only 30 were present in more than 50% of signatures. This core of 30 genes included MITF and many of its targets, indicating that the MITF transcriptional network, which regulates the phenotype switching from the invasive to the proliferative cell state (Hoek et al., 2008; Hoek and Goding, 2010), is captured by the majority of signatures. The signatures lacking MITF and its target genes could include other genes correlated to phenotype switching, thus explaining the poor overlap among them. Shifting from a gene- to a pathway-level did not show a relevant increase in signatures' similarity, with a few exceptions. We hypothesize that this analysis is limited by the incomplete knowledge of gene functions and by the problem of correlated variables in high-throughput studies. Not all genes, indeed, are currently assigned to a pathway and this can negatively impact the outcome of a pathway analysis. A clear example comes from the TCGA classification: despite the keratin and MITF-low subtypes differ in the expression of genes related to melanocyte development, pigmentation signaling and melanin synthesis, neither MITF

nor its target genes were found in TCGA signature. This shows that the genes included in TCGA signatures are correlated to MITF but they are not categorized into melanocyte or pigmentation pathways.

Despite the inconsistencies among the nine gene signatures we observed a high degree of concordance in sample classifications indicating that different subset of genes capture the same underlying biology of melanoma subtypes. Our analytical framework confirmed that the dual invasive/proliferative classification can be further refined in two additional biologically relevant subgroups, in agreement with previous reports (Dugo et al., 2015; Tsoi et al., 2018). Moreover, through our approach we were able to identify a previously undescribed subtype composed of samples that, according to the existing melanoma classifiers, had an undetermined classification. Two of our five subtypes, the undifferentiated and the neural crest-like were characterized by an invasive phenotype, while the other two, the melanoblast-like and the melanocytic subtypes correlated with signatures of proliferation. The newly identified subtype, named transitory/CIN, showed a negative correlation with both invasive and proliferative signatures and had an intermediate expression of markers of both phenotypes. These observations led us hypothesize that this subtype is representative of an intermediate phenotype between the two cell states, hypothesis sustained by the finding that this subtype is located in the middle of the melanocyte differentiation trajectory and reflects a late neural crest/early melanoblast state. Altogether these results suggest that melanoma cells can adopt intermediate states between the invasive and proliferative states; however we cannot exclude that these transcriptional patterns arise from an admixture of invasive and proliferative cells within the same tumour that confounds gene expression signals.

In line with previous findings (Dugo et al., 2015; Tsoi et al., 2018) samples of the undifferentiated subtype showed the lowest expression of MITF, ERBB3 and SOX10 and the highest expression of AXL and EGFR, both at the gene and protein level as identified by our proteomic analysis. The invasive phenotype of the undifferentiated subtype was confirmed by the up-regulation of SERPINE1, ANX1, and CAV1 shown to be involved in invasiveness of melanoma cells (Boudhraa et al., 2014a; Boudhraa et al., 2014b; Lobos-Gonzalez et al., 2013; Lobos-Gonzalez et al., 2014; Verfaillie et al., 2015). In addition, the undifferentiated subtype was characterized by a strong presence of immune components, both assessed by pathway analysis and gene expression deconvolution. This finding supports the notion that melanoma cell states directly influence the tumour

microenvironment and the high immune infiltration of the undifferentiated subtype could reflect a mechanism similar to the inflammation-induced dedifferentiation observed in *in vivo* and *in vitro* models (Landsberg et al., 2012; Holzel and Tuting, 2016). According to this model, MITF represses the pro-inflammatory response by negatively regulating c-JUN expression. Tumours lacking MITF are hyper-responsive to TNF-α due to active c-JUN and elicit a pro-inflammatory response through extensive recruitment of myeloid immune cells within the tumour microenvironment (Riesenberg et al., 2015). The undifferentiated subtype fits well to this model as it showed down-regulation of MITF, up-regulation of JUN and TNF (data not shown), and abundance of pro-tumourigenic macrophages M2. In contrast with this, we observed also a higher abundance of anti-tumour CD8+ T-cells in this subtype, especially in comparison with the neural crest-like. However, the presence alone of cytotoxic T-cells does not indicate an active anti-tumour response. Markers of immune inhibition (CTLA-4, PD-1 and PDL-1) were indeed up-regulated in this subtype, together with signatures of T-cell exhaustion (Tirosh et al., 2016) (data not shown), suggesting that undifferentiated tumours can escape immune surveillance.

The neural crest-like subtype was distinguished by a marked expression of genes involved in extracellular matrix remodeling and migration, in addition to markers of neural crest cells. At the protein level this subtype showed up-regulation of proteins related to invasiveness such as fibronectin (Olbryt et al., 2011; Wouters et al., 2014), the integrin CD49b (Yoshimura et al., 2009; Knutson et al., 1996) and PKC-alpha (Byers et al., 2010) in concert with loss of the epithelial marker E-cadherin. We found also up-regulation of ASNS, a known target of the transcription factor ATF4, a key mediator of the integrated stress response that is essential in the establishment of the invasive phenotype (Falletta et al., 2017).

The proliferative state was divided in the melanoblast-like and the melanocytic subtypes with both expressing markers of pigment cell differentiation and pigment biosynthesis. According to the rheostat MITF model (Carreira et al., 2006; Goding, 2011; Hoek and Goding, 2010), cells expressing MITF can proliferate or differentiate depending on MITF levels and MITF post-translational modifications. Here we found that the melanocytic subtype recapitulated the fully differentiated state of melanocytes and was characterized by over-expression of metabolic genes involved both in glycolysis or oxidative phosphorylation. While these two mechanisms are considered alternative processes in normal cells it has been shown that they can coexist in cancer cells, thus

increasing cell plasticity (Yu et al., 2017). The melanoblast-like subtype represented the preceding step of fully differentiated melanocytes. While genes involved in cell proliferation were up-regulated in both proliferative subtypes, we found a specific up-regulation of genes related to DNA methylation, mRNA transcription and processing indicating a strong transcriptional activity in the melanoblast-like subtype.

The transitory/CIN subtype showed up-regulation of chromatin remodeling, meiotic and cancer-testis genes. The expression of meiotic genes in mitotically dividing somatic cells is a phenomenon called meiomitosis and it can disrupt the mechanisms that control chromosome maintenance and segregation generating chromosomal instability (Lindsey et al., 2013; Rosa et al., 2012; McFarlane and Wakeman, 2017). In line with this, we found a positive enrichment in this subtype of a chromosomal instability signature. However, the relation between chromosomal instability and phenotype switching remains unclear.

5.2 Prognostic relevance of melanoma consensus subtypes

In addition to being biologically relevant, we obtained evidences that consensus subtypes are determinants of prognosis in stage III/IV melanoma patients. The melanocytic and the transitory/CIN subtypes were characterized by the worst prognosis while the undifferentiated and the neural crest-like subtype were associated to a better outcome. Our findings are in agreement with previous reports showing that proliferative and invasive tumours were associated with poor and good outcome, respectively (Jonsson et al., 2010; Rambow et al., 2015; The Cancer Genome Atlas Network, 2015). However, our classification may provide a better stratification of patients. For instance, the proliferative melanoblast-like subtype showed a better outcome compared to the proliferative melanocytic subtype, despite not statistically significant. These findings suggest that the biological differences between two subgroups of the same phenotypic state may affect the outcome of patients. We also showed that the tumour immune infiltration has a critical prognostic role. As previously reported (Jonsson et al., 2010; Bogunovic et al., 2009; The Cancer Genome Atlas Network, 2015), we found that a high level of tumour immune infiltration is positively associated to outcome. In consideration of that, one would expect the undifferentiated subtype to have the best prognosis due to its high level of immune infiltration. However, despite not significantly, the neural crest-like subtype showed a better survival compared to the undifferentiated subtype. This behavior could be explained by the differences of immune cell types observed between

these two subtypes. These results indicate that for prognostic purposes one should consider not only the level of immune infiltration but also its immunosuppressive or pro-immunogenic state and possibly its spatial location within the tumour microenvironment.

5.3 Predictive value of melanoma consensus subtypes

We also focused our attention on the possibility to use our consensus classification to predict response to therapy. Identification of predictive biomarkers is often pursued through pharmacological studies on *in vitro* cell lines. It is thus important that cell lines are representative models of clinical tumours. Here we showed that the five consensus subtypes identified in tumours are recapitulated in melanoma cell lines and present similar gene expression alterations.

Phenotype switching is directly linked to MAPK inhibitors resistance with invasive cells being more drug-tolerant. Depending on its levels, MITF expression was associated to both resistance and sensitivity to MAPK inhibitors (Johannessen et al., 2013; Haq et al., 2013). The consensus classification fits to this model showing that there is no linear dependence between MITF expression and sensitivity to MAPK inhibitors. As we previously reported, $BRAF^{V600E}$ mutant cell lines of the undifferentiated subtype were the most resistant to BRAF inhibitors and the neural crest-like subtype contained both sensitive and resistant cells. Moreover, we observed a different sensitivity between the melanoblast-like and melanocytic subtypes with the latter having higher levels of MITF and higher tolerance to BRAF inhibitors on average. This is in agreement with the rheostat model according to which low levels of MITF induce a slow-cycling invasive phenotype resistant to MAPK inhibitors, intermediate MITF levels enhance proliferation and confer susceptibility to MAPK inhibitors, whereas high MITF expression induces differentiation-mediates senescence and drug resistance (Carreira et al., 2006; Ahn et al., 2017). Unfortunately, no conclusions can be drawn for the transitory/CIN subtype as only one $BRAF^{V600E}$ cell line fell into this subtype. However, this cell line was sensitive to BRAF inhibitors treatment.

Beyond BRAF inhibitors we observed that the meiomitotic transitory/CIN subtype was the most sensitive to YM-155 and BI-2536. YM-155 is a small-molecule inhibitor that blocks survivin (BIRC5) through direct binding to its promoter and induces apoptosis (Nakahara et al., 2007). BI-2536 is another small-molecule inhibitor that inhibits PLK1 (Steegmaier et al., 2007). Both these proteins, which were up-regulated in this subtype, play a key role in regulation of cell cycle and maintenance of chromosomal stability during

mitosis in proliferating cells and their up-regulation promotes cell survival (Colnaghi and Wheatley, 2010; Feng et al., 2009). The melanocytic subtype showed higher sensitivity to a series of HDAC inhibitors, a class of drugs that was shown to repress the expression of MITF (Yokoyama et al., 2008). Taken together these findings show that each consensus subtype is specifically sensitive to different drugs and *in vitro* pharmacogenomic datasets could be useful resources for identification of new druggable targets and alternative therapeutic strategies for tumours refractory to conventional therapies. However, we would like to point out that, despite we found statistically significant associations between consensus subtypes and drug sensitivity data, the range of sensitivity values within each subtype was remarkably wide. This indicates that the consensus classification alone is not an accurate predictor of drug resistance and other determinants of drug tolerance should be considered for a better prediction.

The predictive role of consensus subtypes was totally lost when we analyzed response to BRAF inhibitors in metastatic melanoma patients. According to the observations obtained *in vitro*, one would expect a negative correlation between best overall response and the undifferentiated subtype and a positive correlation with the transitory/CIN or melanoblast-like subtypes. However, assessment of the correlation between subtypes signatures and response to therapy showed inconsistent results within and between datasets, thus preventing us to draw any reliable conclusion. Several biological and technical factors could explain these negative results. First, it is known that phenotypic plasticity is only one of the many mechanisms that confer drug tolerance to melanoma cells. Of consequence, the therapeutic response of each patient depends not only on the transcriptional state of the tumour but also on other genetic alterations that could not be considered in the analyses we performed due to the lack of such information. These additional mechanisms include the composition of the tumour microenvironment and the cross-talk between tumour, stromal and immune cells. The positive enrichment of immune-related genes, and in particular of the interferon-α/γ pathways, in patients showing better response was the most consistent finding across the datasets we analyzed. How the immune systems influences the response to targeted therapy remains unclear, but our observations in conjunction with previous reports (Hugo et al., 2015), highlight the importance of considering the immune status of tumours when predicting targeted drug treatment response. Another factor that could contribute to the discrepancy between *in vitro* and *in vivo* results is related to the measurement of drug

efficacy. In *in vitro* pharmacological studies, drug response can be reported using different statistics, such as AUC or half maximal inhibitory concentration (IC50), but all of them are a direct measure of the amount of cancer cells killed by the treatment. In solid tumours instead, treatment efficacy is clinically measured using the best overall response, defined by the RECIST criteria (Eisenhauer et al., 2009), that is evaluated by taking into account the change in size and number of target (measurable) and non-target (non measurable) lesions at baseline and at regular time-points during treatment. Thus, best overall response is a measure of the response to treatment of patient, not of the single lesion. It is thus possible that the tumour specimen used for pre-treatment genomic profiling is not representative of the measured patient's response. Of consequence, trying to make a parallelism of treatment response between *in vitro* cell lines and clinical tumours may lead to misinterpretations. Finally, we challenge the adequacy of the currently available datasets used to assess the predictive relevance of consensus subtypes. The sample size of these datasets is very limited reducing the statistical power and the robustness of the analyses. They include patients treated with different BRAF inhibitors and in some cases with combinatorial therapy of BRAF and MEK inhibitors. Tumour specimens were collected from metastases located at different anatomical sites that could differ in levels of stromal or immune contamination. All these aspects add further level of variability that can negatively impact downstream analyses. On the basis of these observations, our analyses remain explorative and we believe that any conclusions should be drawn with caution.

5.4 Multi-omics data analysis of melanoma consensus subtypes

The availability of multiple types of high-throughput data for the same subjects allowed us to exclude significant association between the consensus subtypes and mutational status of all genes, mutational load, and mutational signatures. The independence of phenotype-specific gene expression from BRAF mutational status was already reported in the first works of Hoek and colleagues (Hoek et al., 2006; Widmer et al., 2012). The melanoma genomic and transcriptomic subtypes identified by TCGA were completely unrelated (The Cancer Genome Atlas Network, 2015). Through the analysis of copy number data we showed that our consensus subtypes do not differ in terms of ploidy but present varying degree of chromosomal instability, with the transitory/CIN subtype harboring a higher number of focal amplifications and deletions. Despite both aneuploidy and chromosomal instability confers genomic instability and contribute to the

evolution and heterogeneity of cancer cells, they represent related but different mechanisms that are not necessarily correlated (van Jaarsveld and Kops, 2016). We argue that the higher chromosomal instability in this subtype is a consequence of errors that were expanded over different cell divisions occurring in this meiomitotic phenotype. We identified many regions frequently amplified or deleted across patients. These copy number aberrations include amplifications of oncogenes, such as MITF, AKT3, MDM2 and MYC in addition to amplifications of TERT and BRAF and deletions of tumour-suppressor genes such as PTEN, CDKN2A and STK11. Some of these regions were found to be statistically significantly associated to consensus subtypes. For example the region containing CDKN2A was more frequently deleted in the neural crest-like subtype and less often in the undifferentiated subtype. Despite the statistical significance, looking at the frequencies across subtypes of these regions, and the lack of a positive correlation between expression and copy number status of genes representative of consensus subtypes, made us conclude that copy number events could minimally contribute to the definition of the transcriptional state of melanoma cells.

An opposite situation was found when we analyzed DNA methylation and miRNA expression data. Hundreds of CpG sites were differentially methylated across the five consensus subtypes and many of them were negatively correlated with the expression levels of genes included in the consensus subtype signatures. We also found many genes whose expression levels were positively correlated with their methylation status. This suggests that gene methylation is only one of the multiple mechanisms that control expression and its effect could depend on the genomic context of methylation sites. Another possible explanation is that these genes are under a different transcriptional control or the functional effect of methylation could be masked by the effect of other alterations such as mutations or copy number alterations. A large number of miRNAs showed a strong deregulation across consensus subtypes. The most differentially expressed miRNA was hsa-miR-211-5p, known to inhibit melanoma invasion and to be under direct control of MITF via its host gene TRPM1 (Bell et al., 2014; Mazar et al., 2010; Levy et al., 2010). Many miRNAs were found to act in a synergistic way as they targeted the same genes and consequently the same pathways. Other miRNAs instead showed a more specialized regulatory function each targeting specific pathways. Here, we provided a general overview of the molecular alterations underlying these biological entities but a further in deep characterization of genomic abnormalities is necessary to uncover

relevant functional aspects of the disease and drive effective drug development strategies.

6 CONCLUSIONS

Our results showed that melanoma gene expression classifications converged on five fundamental biological entities with distinct transcriptional, epigenomic and microenvironmental features, adding new insights to the knowledge of the disease. Among them, the transitory/CIN melanoma subtype was a previously unrecognized entity characterized by features of genomic instability and expression of cancer-testis antigens. Furthermore, analysis of multi-omics data indicated that phenotypic plasticity of melanoma cells was not determined by acquisition of specific mutations or other genomic alterations, which are irreversible, but it was the result of a transient transcriptional reprogramming under tight epigenetic control. Genomic abnormalities can confer to melanomas an additional degree of heterogeneity that can sustain cancer cell fitness and evolution. Assessing more deeply the molecular features that distinguish these melanoma subtypes would uncover relevant functional aspects of the disease, thus providing candidate biomarkers and driving effective drug development strategies.

7 REFERENCES

Agarwala,S.S. and Kirkwood,J.M. (2000). Temozolomide, a novel alkylating agent with activity in the central nervous system, may improve the treatment of advanced metastatic melanoma. Oncologist 5, 144-151.

Ahn,A., Chatterjee,A., and Eccles,M.R. (2017). The Slow Cycling Phenotype: A Growing Problem for Treatment Resistance in Melanoma. Mol Cancer Ther. 16, 1002-1009.

Alexandrov,L.B., Nik-Zainal,S., Wedge,D.C., Aparicio,S.A., Behjati,S., Biankin,A.V., Bignell,G.R., Bolli,N., Borg,A., Borresen-Dale,A.L., Boyault,S., Burkhardt,B., Butler,A.P., Caldas,C., Davies,H.R., Desmedt,C., Eils,R., Eyfjord,J.E., Foekens,J.A., Greaves,M., Hosoda,F., Hutter,B., Ilicic,T., Imbeaud,S., Imielinski,M., Jager,N., Jones,D.T., Jones,D., Knappskog,S., Kool,M., Lakhani,S.R., Lopez-Otin,C., Martin,S., Munshi,N.C., Nakamura,H., Northcott,P.A., Pajic,M., Papaemmanuil,E., Paradiso,A., Pearson,J.V., Puente,X.S., Raine,K., Ramakrishna,M., Richardson,A.L., Richter,J., Rosenstiel,P., Schlesner,M., Schumacher,T.N., Span,P.N., Teague,J.W., Totoki,Y., Tutt,A.N., Valdes-Mas,R., van Buuren,M.M., van,'., V, Vincent-Salomon,A., Waddell,N., Yates,L.R., Zucman-Rossi,J., Futreal,P.A., McDermott,U., Lichter,P., Meyerson,M., Grimmond,S.M., Siebert,R., Campo,E., Shibata,T., Pfister,S.M., Campbell,P.J., and Stratton,M.R. (2013a). Signatures of mutational processes in human cancer. Nature 500, 415-421.

Alexandrov,L.B., Nik-Zainal,S., Wedge,D.C., Campbell,P.J., and Stratton,M.R. (2013b). Deciphering signatures of mutational processes operative in human cancer. Cell Rep. 3, 246-259.

Ali,Z., Yousaf,N., and Larkin,J. (2013). Melanoma epidemiology, biology and prognosis. EJC. Suppl 11, 81-91.

Almeida,L.G., Sakabe,N.J., deOliveira,A.R., Silva,M.C., Mundstein,A.S., Cohen,T., Chen,Y.T., Chua,R., Gurung,S., Gnjatic,S., Jungbluth,A.A., Caballero,O.L., Bairoch,A., Kiesler,E., White,S.L., Simpson,A.J., Old,L.J., Camargo,A.A., and Vasconcelos,A.T. (2009). CTdatabase: a knowledge-base of high-throughput and curated data on cancer-testis antigens. Nucleic Acids Res 37, D816-D819.

Anders,S., Pyl,P.T., and Huber,W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166-169.

Balch,C.M., Gershenwald,J.E., Soong,S.J., Thompson,J.F., Atkins,M.B., Byrd,D.R., Buzaid,A.C., Cochran,A.J., Coit,D.G., Ding,S., Eggermont,A.M., Flaherty,K.T., Gimotty,P.A., Kirkwood,J.M., McMasters,K.M., Mihm,M.C., Jr., Morton,D.L., Ross,M.I., Sober,A.J., and Sondak,V.K. (2009). Final version of 2009 AJCC melanoma staging and classification. J Clin Oncol 27, 6199-6206.

Balch,C.M., Soong,S.J., Gershenwald,J.E., Thompson,J.F., Coit,D.G., Atkins,M.B., Ding,S., Cochran,A.J., Eggermont,A.M., Flaherty,K.T., Gimotty,P.A., Johnson,T.M., Kirkwood,J.M., Leong,S.P., McMasters,K.M., Mihm,M.C., Jr., Morton,D.L., Ross,M.I., and Sondak,V.K.

(2013). Age as a prognostic factor in patients with localized melanoma and regional metastases. Ann. Surg Oncol 20, 3961-3968.

Balch,C.M., Wilkerson,J.A., Murad,T.M., Soong,S.J., Ingalls,A.L., and Maddox,W.A. (1980). The prognostic significance of ulceration of cutaneous melanoma. Cancer 45, 3012-3017.

Barretina,J., Caponigro,G., Stransky,N., Venkatesan,K., Margolin,A.A., Kim,S., Wilson,C.J., Lehar,J., Kryukov,G.V., Sonkin,D., Reddy,A., Liu,M., Murray,L., Berger,M.F., Monahan,J.E., Morais,P., Meltzer,J., Korejwa,A., Jane-Valbuena,J., Mapa,F.A., Thibault,J., Bric-Furlong,E., Raman,P., Shipway,A., Engels,I.H., Cheng,J., Yu,G.K., Yu,J., Aspesi,P., Jr., de,S.M., Jagtap,K., Jones,M.D., Wang,L., Hatton,C., Palescandolo,E., Gupta,S., Mahan,S., Sougnez,C., Onofrio,R.C., Liefeld,T., MacConaill,L., Winckler,W., Reich,M., Li,N., Mesirov,J.P., Gabriel,S.B., Getz,G., Ardlie,K., Chan,V., Myer,V.E., Weber,B.L., Porter,J., Warmuth,M., Finan,P., Harris,J.L., Meyerson,M., Golub,T.R., Morrissey,M.P., Sellers,W.R., Schlegel,R., and Garraway,L.A. (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature 483, 603-607.

Bell,R.E., Khaled,M., Netanely,D., Schubert,S., Golan,T., Buxbaum,A., Janas,M.M., Postolsky,B., Goldberg,M.S., Shamir,R., and Levy,C. (2014). Transcription factor/microRNA axis blocks melanoma invasion program by miR-211 targeting NUAK1. J Invest Dermatol 134, 441-451.

Benjamini,Y. and Hochberg,Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society 57, 289-300.

Betel,D., Wilson,M., Gabow,A., Marks,D.S., and Sander,C. (2008). The microRNA.org resource: targets and expression. Nucleic Acids Res 36, D149-D153.

Bevona,C., Goggins,W., Quinn,T., Fullerton,J., and Tsao,H. (2003). Cutaneous melanomas associated with nevi. Arch. Dermatol 139, 1620-1624.

Bittner,M., Meltzer,P., Chen,Y., Jiang,Y., Seftor,E., Hendrix,M., Radmacher,M., Simon,R., Yakhini,Z., Ben-Dor,A., Sampas,N., Dougherty,E., Wang,E., Marincola,F., Gooden,C., Lueders,J., Glatfelter,A., Pollock,P., Carpten,J., Gillanders,E., Leja,D., Dietrich,K., Beaudry,C., Berens,M., Alberts,D., and Sondak,V. (2000). Molecular classification of cutaneous malignant melanoma by gene expression profiling. Nature 406, 536-540.

Bogunovic,D., O'Neill,D.W., Belitskaya-Levy,I., Vacic,V., Yu,Y.L., Adams,S., Darvishian,F., Berman,R., Shapiro,R., Pavlick,A.C., Lonardi,S., Zavadil,J., Osman,I., and Bhardwaj,N. (2009). Immune profile and mitotic index of metastatic melanoma lesions enhance clinical staging in predicting patient survival. Proc. Natl Acad. Sci U. S. A 106, 20429-20434.

Boudhraa,Z., Merle,C., Mazzocut,D., Chezal,J.M., Chambon,C., Miot-Noirault,E., Theisen,M., Bouchon,B., and Degoul,F. (2014a). Characterization of pro-invasive mechanisms and N-terminal cleavage of ANXA1 in melanoma. Arch. Dermatol Res 306, 903-914.

Boudhraa,Z., Rondepierre,F., Ouchchane,L., Kintossou,R., Trzeciakiewicz,A., Franck,F., Kanitakis,J., Labeille,B., Joubert-Zakeyh,J., Bouchon,B., Perrot,J.L., Mansard,S., Papon,J., Dechelotte,P., Chezal,J.M., Miot-Noirault,E., Bonnet,M., D'Incan,M., and Degoul,F.

(2014b). Annexin A1 in primary tumors promotes melanoma dissemination. Clin Exp. Metastasis 31, 749-760.

Boyle,G.M., Woods,S.L., Bonazzi,V.F., Stark,M.S., Hacker,E., Aoude,L.G., Dutton-Regester,K., Cook,A.L., Sturm,R.A., and Hayward,N.K. (2011). Melanoma cell invasiveness is regulated by miR-211 suppression of the BRN2 transcription factor. Pigment Cell Melanoma Res 24, 525-537.

Bray,F., Ferlay,J., Soerjomataram,I., Siegel,R.L., Torre,L.A., and Jemal,A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin.

Brenner,M. and Hearing,V.J. (2008). The protective role of melanin against UV damage in human skin. Photochem. Photobiol. 84, 539-549.

Breslow,A. (1970). Thickness, cross-sectional areas and depth of invasion in the prognosis of cutaneous melanoma. Ann. Surg 172, 902-908.

Brunet,J.F., Denizot,F., Luciani,M.F., Roux-Dosseto,M., Suzan,M., Mattei,M.G., and Golstein,P. (1987). A new member of the immunoglobulin superfamily--CTLA-4. Nature 328, 267-270.

Byers,H.R., Boissel,S.J., Tu,C., and Park,H.Y. (2010). RNAi-mediated knockdown of protein kinase C-alpha inhibits cell migration in MM-RU human metastatic melanoma cell line. Melanoma Res 20, 171-178.

Carreira,S., Goodall,J., Denat,L., Rodriguez,M., Nuciforo,P., Hoek,K.S., Testori,A., Larue,L., and Goding,C.R. (2006). Mitf regulation of Dia1 controls melanoma proliferation and invasiveness. Genes Dev. 20, 3426-3439.

Carter,S.L., Cibulskis,K., Helman,E., McKenna,A., Shen,H., Zack,T., Laird,P.W., Onofrio,R.C., Winckler,W., Weir,B.A., Beroukhim,R., Pellman,D., Levine,D.A., Lander,E.S., Meyerson,M., and Getz,G. (2012). Absolute quantification of somatic DNA alterations in human cancer. Nat Biotechnol. 30, 413-421.

Carter,S.L., Eklund,A.C., Kohane,I.S., Harris,L.N., and Szallasi,Z. (2006). A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. Nat Genet. 38, 1043-1048.

Cavalcante,R.G. and Sartor,M.A. (2017). annotatr: genomic regions in context. Bioinformatics 33, 2381-2383.

Chao,C., Martin,R.C., Ross,M.I., Reintgen,D.S., Edwards,M.J., Noyes,R.D., Hagendoorn,L.J., Stromberg,A.J., and McMasters,K.M. (2004). Correlation between prognostic factors and increasing age in melanoma. Ann. Surg Oncol 11, 259-264.

Chapman,P.B., Hauschild,A., Robert,C., Haanen,J.B., Ascierto,P., Larkin,J., Dummer,R., Garbe,C., Testori,A., Maio,M., Hogg,D., Lorigan,P., Lebbe,C., Jouary,T., Schadendorf,D., Ribas,A., O'Day,S.J., Sosman,J.A., Kirkwood,J.M., Eggermont,A.M., Dreno,B., Nolop,K., Li,J., Nelson,B., Hou,J., Lee,R.J., Flaherty,K.T., and McArthur,G.A. (2011). Improved survival with vemurafenib in melanoma with BRAF V600E mutation. N. Engl. J Med. 364, 2507-2516.

Charoentong,P., Finotello,F., Angelova,M., Mayer,C., Efremova,M., Rieder,D., Hackl,H., and Trajanoski,Z. (2017). Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. Cell Rep. 18, 248-262.

Chen,H., Weng,Q.Y., and Fisher,D.E. (2014). UV signaling pathways within the skin. J Invest Dermatol 134, 2080-2085.

Colnaghi,R. and Wheatley,S.P. (2010). Liaisons between survivin and Plk1 during cell division and cell death. J Biol Chem. 285, 22592-22604.

Conway,C., Mitra,A., Jewell,R., Randerson-Moor,J., Lobo,S., Nsengimana,J., Edward,S., Sanders,D.S., Cook,M., Powell,B., Boon,A., Elliott,F., de,K.F., Knowles,M.A., Bishop,D.T., and Newton-Bishop,J. (2009). Gene expression profiling of paraffin-embedded primary melanoma using the DASL assay identifies increased osteopontin expression as predictive of reduced relapse-free survival. Clin Cancer Res 15, 6939-6946.

Cust,A.E., Armstrong,B.K., Goumas,C., Jenkins,M.A., Schmid,H., Hopper,J.L., Kefford,R.F., Giles,G.G., Aitken,J.F., and Mann,G.J. (2011). Sunbed use during adolescence and early adulthood is associated with increased risk of early-onset melanoma. Int J Cancer 128, 2425-2435.

Davies,H., Bignell,G.R., Cox,C., Stephens,P., Edkins,S., Clegg,S., Teague,J., Woffendin,H., Garnett,M.J., Bottomley,W., Davis,N., Dicks,E., Ewing,R., Floyd,Y., Gray,K., Hall,S., Hawes,R., Hughes,J., Kosmidou,V., Menzies,A., Mould,C., Parker,A., Stevens,C., Watt,S., Hooper,S., Wilson,R., Jayatilake,H., Gusterson,B.A., Cooper,C., Shipley,J., Hargrave,D., Pritchard-Jones,K., Maitland,N., Chenevix-Trench,G., Riggins,G.J., Bigner,D.D., Palmieri,G., Cossu,A., Flanagan,A., Nicholson,A., Ho,J.W., Leung,S.Y., Yuen,S.T., Weber,B.L., Seigler,H.F., Darrow,T.L., Paterson,H., Marais,R., Marshall,C.J., Wooster,R., Stratton,M.R., and Futreal,P.A. (2002). Mutations of the BRAF gene in human cancer. Nature 417, 949-954.

DiFronzo,L.A., Wanek,L.A., Elashoff,R., and Morton,D.L. (1999). Increased incidence of second primary melanoma in patients with a previous cutaneous melanoma. Ann. Surg Oncol 6, 705-711.

Du,P., Kibbe,W.A., and Lin,S.M. (2008). lumi: a pipeline for processing Illumina microarray. Bioinformatics 24, 1547-1548.

Dugo,M., Nicolini,G., Tragni,G., Bersani,I., Tomassetti,A., Colonna,V., Del,V.M., De,B.F., Canevari,S., Anichini,A., and Sensi,M. (2015). A melanoma subtype with intrinsic resistance to BRAF inhibition identified by receptor tyrosine kinases gene-driven classification. Oncotarget 6, 5118-5133.

Eggermont,A.M. and Kirkwood,J.M. (2004). Re-evaluating the role of dacarbazine in metastatic melanoma: what have we learned in 30 years? Eur J Cancer 40, 1825-1836.

Eichhoff,O.M., Zipser,M.C., Xu,M., Weeraratna,A.T., Mihic,D., Dummer,R., and Hoek,K.S. (2010). The immunohistochemistry of invasive and proliferative phenotype switching in melanoma: a case report. Melanoma Res 20, 349-355.

Eisenhauer,E.A., Therasse,P., Bogaerts,J., Schwartz,L.H., Sargent,D., Ford,R., Dancey,J., Arbuck,S., Gwyther,S., Mooney,M., Rubinstein,L., Shankar,L., Dodd,L., Kaplan,R., Lacombe,D., and Verweij,J. (2009). New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer 45, 228-247.

Elwood,J.M. and Jopson,J. (1997). Melanoma and sun exposure: an overview of published studies. Int J Cancer 73, 198-203.

Fadaki,N., Li,R., Parrett,B., Sanders,G., Thummala,S., Martineau,L., Cardona-Huerta,S., Miranda,S., Cheng,S.T., Miller,J.R., III, Singer,M., Cleaver,J.E., Kashani-Sabet,M., and Leong,S.P. (2013). Is head and neck melanoma different from trunk and extremity melanomas with respect to sentinel lymph node status and clinical outcome? Ann. Surg Oncol 20, 3089-3097.

Falletta,P., Sanchez-Del-Campo,L., Chauhan,J., Effern,M., Kenyon,A., Kershaw,C.J., Siddaway,R., Lisle,R., Freter,R., Daniels,M.J., Lu,X., Tuting,T., Middleton,M., Buffa,F.M., Willis,A.E., Pavitt,G., Ronai,Z.A., Sauka-Spengler,T., Holzel,M., and Goding,C.R. (2017). Translation reprogramming is an evolutionarily conserved driver of phenotypic plasticity and therapeutic resistance in melanoma. Genes Dev. 31, 18-33.

Feng,Y.B., Lin,D.C., Shi,Z.Z., Wang,X.C., Shen,X.M., Zhang,Y., Du,X.L., Luo,M.L., Xu,X., Han,Y.L., Cai,Y., Zhang,Z.Q., Zhan,Q.M., and Wang,M.R. (2009). Overexpression of PLK1 is associated with poor survival by inhibiting apoptosis via enhancement of survivin level in esophageal squamous cell carcinoma. Int J Cancer 124, 578-588.

Fioranelli,M., Roccia,M.G., Pastore,C., Aracena,C.J., and Lotti,T. (2017). Completion dissection or observation for sentinel-node metastasis in melanoma. Dermatol Ther. 30.

Flaherty,K.T., Robert,C., Hersey,P., Nathan,P., Garbe,C., Milhem,M., Demidov,L.V., Hassel,J.C., Rutkowski,P., Mohr,P., Dummer,R., Trefzer,U., Larkin,J.M., Utikal,J., Dreno,B., Nyakas,M., Middleton,M.R., Becker,J.C., Casey,M., Sherman,L.J., Wu,F.S., Ouellet,D., Martin,A.M., Patel,K., and Schadendorf,D. (2012). Improved survival with MEK inhibition in BRAF-mutated melanoma. N. Engl. J Med. 367, 107-114.

Gershenwald,J.E., Scolyer,R.A., Hess,K.R., Sondak,V.K., Long,G.V., Ross,M.I., Lazar,A.J., Faries,M.B., Kirkwood,J.M., McArthur,G.A., Haydu,L.E., Eggermont,A.M.M., Flaherty,K.T., Balch,C.M., and Thompson,J.F. (2017). Melanoma staging: Evidence-based changes in the American Joint Committee on Cancer eighth edition cancer staging manual. CA Cancer J Clin 67, 472-492.

Gilchrest,B.A., Eller,M.S., Geller,A.C., and Yaar,M. (1999). The pathogenesis of melanoma induced by ultraviolet radiation. N. Engl. J Med. 340, 1341-1348.

Goding,C.R. (2011). Commentary. A picture of Mitf in melanoma immortality. Oncogene 30, 2304-2306.

Goldstein,A.M. and Tucker,M.A. (2001). Genetic epidemiology of cutaneous melanoma: a global perspective. Arch. Dermatol 137, 1493-1496.

Golub,T.R., Slonim,D.K., Tamayo,P., Huard,C., Gaasenbeek,M., Mesirov,J.P., Coller,H., Loh,M.L., Downing,J.R., Caligiuri,M.A., Bloomfield,C.D., and Lander,E.S. (1999). Molecular

classification of cancer: class discovery and class prediction by gene expression monitoring. Science 286, 531-537.

Grob,J.J., Gouvernet,J., Aymar,D., Mostaque,A., Romano,M.H., Collet,A.M., Noe,M.C., Diconstanzo,M.P., and Bonerandi,J.J. (1990). Count of benign melanocytic nevi as a major indicator of risk for nonfamilial nodular and superficial spreading melanoma. Cancer 66, 387-395.

Gruis,N.A., van,d., V, Sandkuijl,L.A., Prins,D.E., Weaver-Feldhaus,J., Kamb,A., Bergman,W., and Frants,R.R. (1995). Homozygotes for CDKN2 (p16) germline mutation in Dutch familial melanoma kindreds. Nat Genet. 10, 351-353.

Hanzelmann,S., Castelo,R., and Guinney,J. (2013). GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics 14, 7.

Haq,R., Yokoyama,S., Hawryluk,E.B., Jonsson,G.B., Frederick,D.T., McHenry,K., Porter,D., Tran,T.N., Love,K.T., Langer,R., Anderson,D.G., Garraway,L.A., Duncan,L.M., Morton,D.L., Hoon,D.S., Wargo,J.A., Song,J.S., and Fisher,D.E. (2013). BCL2A1 is a lineage-specific antiapoptotic melanoma oncogene that confers resistance to BRAF inhibition. Proc. Natl. Acad. Sci U. S. A 110, 4321-4326.

Haqq,C., Nosrati,M., Sudilovsky,D., Crothers,J., Khodabakhsh,D., Pulliam,B.L., Federman,S., Miller,J.R., III, Allen,R.E., Singer,M.I., Leong,S.P., Ljung,B.M., Sagebiel,R.W., and Kashani-Sabet,M. (2005). The gene expression signatures of melanoma progression. Proc. Natl. Acad. Sci U. S. A 102, 6092-6097.

Hauschild,A., Grob,J.J., Demidov,L.V., Jouary,T., Gutzmer,R., Millward,M., Rutkowski,P., Blank,C.U., Miller,W.H., Jr., Kaempgen,E., Martin-Algarra,S., Karaszewska,B., Mauch,C., Chiarion-Sileni,V., Martin,A.M., Swann,S., Haney,P., Mirakhur,B., Guckert,M.E., Goodman,V., and Chapman,P.B. (2012). Dabrafenib in BRAF-mutated metastatic melanoma: a multicentre, open-label, phase 3 randomised controlled trial. Lancet 380, 358-365.

Hodi,F.S., O'Day,S.J., McDermott,D.F., Weber,R.W., Sosman,J.A., Haanen,J.B., Gonzalez,R., Robert,C., Schadendorf,D., Hassel,J.C., Akerley,W., van den Eertwegh,A.J., Lutzky,J., Lorigan,P., Vaubel,J.M., Linette,G.P., Hogg,D., Ottensmeier,C.H., Lebbe,C., Peschel,C., Quirt,I., Clark,J.I., Wolchok,J.D., Weber,J.S., Tian,J., Yellin,M.J., Nichol,G.M., Hoos,A., and Urba,W.J. (2010). Improved survival with ipilimumab in patients with metastatic melanoma. N. Engl. J Med. 363, 711-723.

Hoek,K.S., Eichhoff,O.M., Schlegel,N.C., Dobbeling,U., Kobert,N., Schaerer,L., Hemmi,S., and Dummer,R. (2008). In vivo switching of human melanoma cells between proliferative and invasive states. Cancer Res 68, 650-656.

Hoek,K.S. and Goding,C.R. (2010). Cancer stem cells versus phenotype-switching in melanoma. Pigment Cell Melanoma Res 23, 746-759.

Hoek,K.S., Schlegel,N.C., Brafford,P., Sucker,A., Ugurel,S., Kumar,R., Weber,B.L., Nathanson,K.L., Phillips,D.J., Herlyn,M., Schadendorf,D., and Dummer,R. (2006). Metastatic potential of melanomas defined by specific gene expression profiles with no BRAF signature. Pigment Cell Res 19, 290-302.

Holly,E.A., Kelly,J.W., Shpall,S.N., and Chiu,S.H. (1987). Number of melanocytic nevi as a major risk factor for malignant melanoma. J Am Acad. Dermatol 17, 459-468.

Holman,C.D. and Armstrong,B.K. (1984). Cutaneous malignant melanoma and indicators of total accumulated exposure to the sun: an analysis separating histogenetic types. J Natl. Cancer Inst. 73, 75-82.

Holzel,M. and Tuting,T. (2016). Inflammation-Induced Plasticity in Melanoma Therapy and Metastasis. Trends Immunol. 37, 364-374.

Hoshida,Y. (2010). Nearest template prediction: a single-sample-based flexible class prediction with confidence assessment. PLoS One 5, e15543.

Hsu,M.Y., Wheelock,M.J., Johnson,K.R., and Herlyn,M. (1996). Shifts in cadherin profiles between human normal melanocytes and melanomas. J Investig. Dermatol Symp. Proc. 1, 188-194.

Hugo,W., Shi,H., Sun,L., Piva,M., Song,C., Kong,X., Moriceau,G., Hong,A., Dahlman,K.B., Johnson,D.B., Sosman,J.A., Ribas,A., and Lo,R.S. (2015). Non-genomic and Immune Evolution of Melanoma Acquiring MAPKi Resistance. Cell 162, 1271-1285.

Hugo,W., Zaretsky,J.M., Sun,L., Song,C., Moreno,B.H., Hu-Lieskovan,S., Berent-Maoz,B., Pang,J., Chmielowski,B., Cherry,G., Seja,E., Lomeli,S., Kong,X., Kelley,M.C., Sosman,J.A., Johnson,D.B., Ribas,A., and Lo,R.S. (2016). Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma. Cell 165, 35-44.

Irizarry,R.A., Hobbs,B., Collin,F., Beazer-Barclay,Y.D., Antonellis,K.J., Scherf,U., and Speed,T.P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 4, 249-264.

Johannessen,C.M., Boehm,J.S., Kim,S.Y., Thomas,S.R., Wardwell,L., Johnson,L.A., Emery,C.M., Stransky,N., Cogdill,A.P., Barretina,J., Caponigro,G., Hieronymus,H., Murray,R.R., Salehi-Ashtiani,K., Hill,D.E., Vidal,M., Zhao,J.J., Yang,X., Alkan,O., Kim,S., Harris,J.L., Wilson,C.J., Myer,V.E., Finan,P.M., Root,D.E., Roberts,T.M., Golub,T., Flaherty,K.T., Dummer,R., Weber,B.L., Sellers,W.R., Schlegel,R., Wargo,J.A., Hahn,W.C., and Garraway,L.A. (2010). COT drives resistance to RAF inhibition through MAP kinase pathway reactivation. Nature 468, 968-972.

Johannessen,C.M., Johnson,L.A., Piccioni,F., Townes,A., Frederick,D.T., Donahue,M.K., Narayan,R., Flaherty,K.T., Wargo,J.A., Root,D.E., and Garraway,L.A. (2013). A melanocyte lineage program confers resistance to MAP kinase pathway inhibition. Nature 504, 138-142.

John,T., Black,M.A., Toro,T.T., Leader,D., Gedye,C.A., Davis,I.D., Guilford,P.J., and Cebon,J.S. (2008). Predicting clinical outcome through molecular profiling in stage III melanoma. Clin Cancer Res 14, 5173-5180.

Johnson,W.E., Li,C., and Rabinovic,A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 8, 118-127.

Jonsson,G., Busch,C., Knappskog,S., Geisler,J., Miletic,H., Ringner,M., Lillehaug,J.R., Borg,A., and Lonning,P.E. (2010). Gene expression profiling-based identification of

molecular subtypes in stage IV melanomas with different clinical outcome. Clin Cancer Res 16, 3356-3367.

Joosse,A., Collette,S., Suciu,S., Nijsten,T., Lejeune,F., Kleeberg,U.R., Coebergh,J.W., Eggermont,A.M., and de,V.E. (2012). Superior outcome of women with stage I/II cutaneous melanoma: pooled analysis of four European Organisation for Research and Treatment of Cancer phase III trials. J Clin Oncol 30, 2240-2247.

Joosse,A., Collette,S., Suciu,S., Nijsten,T., Patel,P.M., Keilholz,U., Eggermont,A.M., Coebergh,J.W., and de,V.E. (2013). Sex is an independent prognostic indicator for survival and relapse/progression-free survival in metastasized stage III to IV melanoma: a pooled analysis of five European organisation for research and treatment of cancer randomized controlled trials. J Clin Oncol 31, 2337-2346.

Kabbarah,O., Nogueira,C., Feng,B., Nazarian,R.M., Bosenberg,M., Wu,M., Scott,K.L., Kwong,L.N., Xiao,Y., Cordon-Cardo,C., Granter,S.R., Ramaswamy,S., Golub,T., Duncan,L.M., Wagner,S.N., Brennan,C., and Chin,L. (2010). Integrative genome comparison of primary and metastatic melanomas. PLoS One 5, e10770.

Kanetsky,P.A., Rebbeck,T.R., Hummer,A.J., Panossian,S., Armstrong,B.K., Kricker,A., Marrett,L.D., Millikan,R.C., Gruber,S.B., Culver,H.A., Zanetti,R., Gallagher,R.P., Dwyer,T., Busam,K., From,L., Mujumdar,U., Wilcox,H., Begg,C.B., and Berwick,M. (2006). Population-based study of natural variation in the melanocortin-1 receptor gene and melanoma. Cancer Res 66, 9330-9337.

Kertesz,M., Iovino,N., Unnerstall,U., Gaul,U., and Segal,E. (2007). The role of site accessibility in microRNA target recognition. Nat Genet. 39, 1278-1284.

Kim,C., Lee,C.W., Kovacic,L., Shah,A., Klasa,R., and Savage,K.J. (2010). Long-term survival in patients with metastatic melanoma treated with DTIC or temozolomide. Oncologist 15, 765-771.

Kingham,T.P., Panageas,K.S., Ariyan,C.E., Busam,K.J., Brady,M.S., and Coit,D.G. (2010). Outcome of patients with a positive sentinel lymph node who do not undergo completion lymphadenectomy. Ann. Surg Oncol 17, 514-520.

Knutson,J.R., Iida,J., Fields,G.B., and McCarthy,J.B. (1996). CD44/chondroitin sulfate proteoglycan and alpha 2 beta 1 integrin mediate human melanoma cell migration on type IV collagen and invasion of basement membranes. Mol Biol Cell 7, 383-396.

Konieczkowski,D.J., Johannessen,C.M., Abudayyeh,O., Kim,J.W., Cooper,Z.A., Piris,A., Frederick,D.T., Barzily-Rokni,M., Straussman,R., Haq,R., Fisher,D.E., Mesirov,J.P., Hahn,W.C., Flaherty,K.T., Wargo,J.A., Tamayo,P., and Garraway,L.A. (2014). A melanoma cell state distinction influences sensitivity to MAPK pathway inhibitors. Cancer Discov. 4, 816-827.

Kwong,L.N., Boland,G.M., Frederick,D.T., Helms,T.L., Akid,A.T., Miller,J.P., Jiang,S., Cooper,Z.A., Song,X., Seth,S., Kamara,J., Protopopov,A., Mills,G.B., Flaherty,K.T., Wargo,J.A., and Chin,L. (2015). Co-clinical assessment identifies patterns of BRAF inhibitor resistance in melanoma. J Clin Invest 125, 1459-1470.

Landsberg,J., Kohlmeyer,J., Renn,M., Bald,T., Rogava,M., Cron,M., Fatho,M., Lennerz,V., Wolfel,T., Holzel,M., and Tuting,T. (2012). Melanomas resist T-cell therapy through inflammation-induced reversible dedifferentiation. Nature 490, 412-416.

Langfelder,P. and Horvath,S. (2008). WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9, 559.

Larkin,J., Ascierto,P.A., Dreno,B., Atkinson,V., Liszkay,G., Maio,M., Mandala,M., Demidov,L., Stroyakovskiy,D., Thomas,L., de,l.C.-M., Dutriaux,C., Garbe,C., Sovak,M.A., Chang,I., Choong,N., Hack,S.P., McArthur,G.A., and Ribas,A. (2014). Combined vemurafenib and cobimetinib in BRAF-mutated melanoma. N. Engl. J Med. 371, 1867-1876.

Lauss,M., Nsengimana,J., Staaf,J., Newton-Bishop,J., and Jonsson,G. (2016). Consensus of Melanoma Gene Expression Subtypes Converges on Biological Entities. J Invest Dermatol 136, 2502-2505.

Law,C.W., Chen,Y., Shi,W., and Smyth,G.K. (2014). voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol 15, R29.

Leiter,U., Stadler,R., Mauch,C., Hohenberger,W., Brockmeyer,N., Berking,C., Sunderkotter,C., Kaatz,M., Schulte,K.W., Lehmann,P., Vogt,T., Ulrich,J., Herbst,R., Gehring,W., Simon,J.C., Keim,U., Martus,P., and Garbe,C. (2016). Complete lymph node dissection versus no dissection in patients with sentinel lymph node biopsy positive melanoma (DeCOG-SLT): a multicentre, randomised, phase 3 trial. Lancet Oncol 17, 757-767.

Levi,F., Randimbison,L., Te,V.C., and La,V.C. (2005). High constant incidence rates of second cutaneous melanomas. Int J Cancer 117, 877-879.

Levy,C., Khaled,M., Iliopoulos,D., Janas,M.M., Schubert,S., Pinner,S., Chen,P.H., Li,S., Fletcher,A.L., Yokoyama,S., Scott,K.L., Garraway,L.A., Song,J.S., Granter,S.R., Turley,S.J., Fisher,D.E., and Novina,C.D. (2010). Intronic miR-211 assumes the tumor suppressive function of its host gene in melanoma. Mol Cell 40, 841-849.

Lewis,B.P., Burge,C.B., and Bartel,D.P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120, 15-20.

Li,H., Handsaker,B., Wysoker,A., Fennell,T., Ruan,J., Homer,N., Marth,G., Abecasis,G., and Durbin,R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics. 25, 2078-2079.

Liberzon,A., Birger,C., Thorvaldsdottir,H., Ghandi,M., Mesirov,J.P., and Tamayo,P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. 1, 417-425.

Liberzon,A., Subramanian,A., Pinchback,R., Thorvaldsdottir,H., Tamayo,P., and Mesirov,J.P. (2011). Molecular signatures database (MSigDB) 3.0. Bioinformatics 27, 1739-1740.

Lindsey,S.F., Byrnes,D.M., Eller,M.S., Rosa,A.M., Dabas,N., Escandon,J., and Grichnik,J.M. (2013). Potential role of meiosis proteins in melanoma chromosomal instability. J Skin Cancer 2013, 190109.

Lito,P., Pratilas,C.A., Joseph,E.W., Tadi,M., Halilovic,E., Zubrowski,M., Huang,A., Wong,W.L., Callahan,M.K., Merghoub,T., Wolchok,J.D., de,S.E., Chandarlapaty,S., Poulikakos,P.I., Fagin,J.A., and Rosen,N. (2012). Relief of profound feedback inhibition of mitogenic signaling by RAF inhibitors attenuates their activity in BRAFV600E melanomas. Cancer Cell 22, 668-682.

Liu,J., Lichtenberg,T., Hoadley,K.A., Poisson,L.M., Lazar,A.J., Cherniack,A.D., Kovatich,A.J., Benz,C.C., Levine,D.A., Lee,A.V., Omberg,L., Wolf,D.M., Shriver,C.D., Thorsson,V., and Hu,H. (2018). An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. Cell 173, 400-416.

Lobos-Gonzalez,L., Aguilar,L., Diaz,J., Diaz,N., Urra,H., Torres,V.A., Silva,V., Fitzpatrick,C., Lladser,A., Hoek,K.S., Leyton,L., and Quest,A.F. (2013). E-cadherin determines Caveolin-1 tumor suppression or metastasis enhancing function in melanoma cells. Pigment Cell Melanoma Res 26, 555-570.

Lobos-Gonzalez,L., guilar-Guzman,L., Fernandez,J.G., Munoz,N., Hossain,M., Bieneck,S., Silva,V., Burzio,V., Sviderskaya,E.V., Bennett,D.C., Leyton,L., and Quest,A.F. (2014). Caveolin-1 is a risk factor for postsurgery metastasis in preclinical melanoma models. Melanoma Res 24, 108-119.

Mann,G.J., Pupo,G.M., Campain,A.E., Carter,C.D., Schramm,S.J., Pianova,S., Gerega,S.K., De,S.C., Lai,K., Wilmott,J.S., Synnott,M., Hersey,P., Kefford,R.F., Thompson,J.F., Yang,Y.H., and Scolyer,R.A. (2013). BRAF mutation, NRAS mutation, and the absence of an immune-related expressed gene profile predict poor outcome in patients with stage III melanoma. J Invest Dermatol 133, 509-517.

Martinez,S.R. and Young,S.E. (2008). A rational surgical approach to the treatment of distant melanoma metastases. Cancer Treat. Rev. 34, 614-620.

Matthews,N.H., Li,W.Q., Qureshi,A.A., Weinstock,M.A., and Cho,E. (2017). Epidemiology of Melanoma. In Cutaneous melanoma: Etiology and Therapy, W.H.Ward and J.M.Farma, eds. (Brisbane: Codon Publications).

Mayakonda,A. and Koeffler,H.P. (2016). Maftools: Efficient analysis, visualization and summarization of MAF files from large-scale cohort based cancer studies. bioRxiv.

Mazar,J., DeYoung,K., Khaitan,D., Meister,E., Almodovar,A., Goydos,J., Ray,A., and Perera,R.J. (2010). The regulation of miRNA-211 expression and its role in melanoma cell invasiveness. PLoS One 5, e13779.

McCall,M.N., Bolstad,B.M., and Irizarry,R.A. (2010). Frozen robust multiarray analysis (fRMA). Biostatistics 11, 242-253.

McFarlane,R.J. and Wakeman,J.A. (2017). Meiosis-like Functions in Oncogenesis: A New View of Cancer. Cancer Res 77, 5712-5716.

McGovern,V.J., Shaw,H.M., Milton,G.W., and McCarthy,W.H. (1982). Ulceration and prognosis in cutaneous malignant melanoma. Histopathology 6, 399-407.

McGranahan,N. and Swanton,C. (2017). Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. Cell 168, 613-628.

Merico,D., Isserlin,R., Stueker,O., Emili,A., and Bader,G.D. (2010). Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. PLoS One 5, e13984.

Merimsky,O. and Inbar,M. (1998). Cigarette smoking and skin cancer. Clin Dermatol 16, 585-588.

Mermel,C.H., Schumacher,S.E., Hill,B., Meyerson,M.L., Beroukhim,R., and Getz,G. (2011). GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol 12, R41.

Mica,Y., Lee,G., Chambers,S.M., Tomishima,M.J., and Studer,L. (2013). Modeling neural crest induction, melanocyte specification, and disease-related pigmentation defects in hESCs and patient-specific iPSCs. Cell Rep. 3, 1140-1152.

Michaloglou,C., Vredeveld,L.C., Soengas,M.S., Denoyelle,C., Kuilman,T., van der Horst,C.M., Majoor,D.M., Shay,J.W., Mooi,W.J., and Peeper,D.S. (2005). BRAFE600-associated senescence-like cell cycle arrest of human naevi. Nature 436, 720-724.

Middleton,M.R., Grob,J.J., Aaronson,N., Fierlbeck,G., Tilgen,W., Seiter,S., Gore,M., Aamdal,S., Cebon,J., Coates,A., Dreno,B., Henz,M., Schadendorf,D., Kapp,A., Weiss,J., Fraass,U., Statkevich,P., Muller,M., and Thatcher,N. (2000). Randomized phase III study of temozolomide versus dacarbazine in the treatment of patients with advanced metastatic malignant melanoma. J Clin Oncol 18, 158-166.

Miller,A.J. and Mihm,M.C., Jr. (2006). Melanoma. N. Engl. J Med. 355, 51-65.

Miller,J.A., Cai,C., Langfelder,P., Geschwind,D.H., Kurian,S.M., Salomon,D.R., and Horvath,S. (2011). Strategies for aggregating gene expression data: the collapseRows R function. BMC Bioinformatics 12, 322.

Miranda,K.C., Huynh,T., Tay,Y., Ang,Y.S., Tam,W.L., Thomson,A.M., Lim,B., and Rigoutsos,I. (2006). A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. Cell 126, 1203-1217.

Monti,S., Tamayo,P., Mesirov,J.P., and Golub,T.R. (2003). Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. Machine Learning 52, 91-118.

Morgan,A.M., Lo,J., and Fisher,D.E. (2013). How does pheomelanin synthesis contribute to melanomagenesis?: Two distinct mechanisms could explain the carcinogenicity of pheomelanin synthesis. Bioessays 35, 672-676.

Muller,J., Krijgsman,O., Tsoi,J., Robert,L., Hugo,W., Song,C., Kong,X., Possik,P.A., Cornelissen-Steijger,P.D., Geukes Foppen,M.H., Kemper,K., Goding,C.R., McDermott,U., Blank,C., Haanen,J., Graeber,T.G., Ribas,A., Lo,R.S., and Peeper,D.S. (2014). Low MITF/AXL

ratio predicts early resistance to multiple targeted drugs in melanoma. Nat Commun. 5, 5712.

Nakahara,T., Kita,A., Yamanaka,K., Mori,M., Amino,N., Takeuchi,M., Tominaga,F., Hatakeyama,S., Kinoyama,I., Matsuhisa,A., Kudoh,M., and Sasamata,M. (2007). YM155, a novel small-molecule survivin suppressant, induces regression of established human hormone-refractory prostate tumor xenografts. Cancer Res 67, 8014-8021.

Nazarian,R., Shi,H., Wang,Q., Kong,X., Koya,R.C., Lee,H., Chen,Z., Lee,M.K., Attar,N., Sazegar,H., Chodon,T., Nelson,S.F., McArthur,G., Sosman,J.A., Ribas,A., and Lo,R.S. (2010). Melanomas acquire resistance to B-RAF(V600E) inhibition by RTK or N-RAS upregulation. Nature 468, 973-977.

Newman,A.M., Liu,C.L., Green,M.R., Gentles,A.J., Feng,W., Xu,Y., Hoang,C.D., Diehn,M., and Alizadeh,A.A. (2015). Robust enumeration of cell subsets from tissue expression profiles. Nat. Methods 12, 453-457.

Noone, A. M., Howlader, N., Krapcho, M., Miller, D., Brest, A., Yu, M., Ruhl, J., Tatalovich, Z., Mariotto, A., Lewis, D. R., Chen, H. S., Feuer, E. J., and Cronin, K. A. eds. SEER Cancer Statistics Review, 1975-2015, National Cancer Institute. Bethesda, MD, https://seer.cancer.gov/csr/1975_2015/, based on November 2017 SEER data submission, posted to the SEER web site, April 2018. 1-11-2017.
Ref Type: Generic

Okazaki,T. and Honjo,T. (2006). The PD-1-PD-L pathway in immunological tolerance. Trends Immunol. 27, 195-201.

Olbryt,M., Habryka,A., Tyszkiewicz,T., Rusin,A., Cichon,T., Jarzab,M., and Krawczyk,Z. (2011). Melanoma-associated genes, MXI1, FN1, and NME1, are hypoxia responsive in murine and human melanoma cells. Melanoma Res 21, 417-425.

Paraiso,K.H., Xiang,Y., Rebecca,V.W., Abel,E.V., Chen,Y.A., Munko,A.C., Wood,E., Fedorenko,I.V., Sondak,V.K., Anderson,A.R., Ribas,A., Palma,M.D., Nathanson,K.L., Koomen,J.M., Messina,J.L., and Smalley,K.S. (2011). PTEN loss confers BRAF inhibitor resistance to melanoma cells through the suppression of BIM expression. Cancer Res 71, 2750-2760.

Patel,P.M., Suciu,S., Mortier,L., Kruit,W.H., Robert,C., Schadendorf,D., Trefzer,U., Punt,C.J., Dummer,R., Davidson,N., Becker,J., Conry,R., Thompson,J.A., Hwu,W.J., Engelen,K., Agarwala,S.S., Keilholz,U., Eggermont,A.M., and Spatz,A. (2011). Extended schedule, escalated dose temozolomide versus dacarbazine in stage IV melanoma: final results of a randomised phase III study (EORTC 18032). Eur J Cancer 47, 1476-1483.

Penna,E., Orso,F., Cimino,D., Vercellino,I., Grassi,E., Quaglino,E., Turco,E., and Taverna,D. (2013). miR-214 coordinates melanoma progression by upregulating ALCAM through TFAP2 and miR-148b downmodulation. Cancer Res 73, 4098-4111.

Peters,T.J., Buckley,M.J., Statham,A.L., Pidsley,R., Samaras,K., Lord,V., Clark,S.J., and Molloy,P.L. (2015). De novo identification of differentially methylated regions in the human genome. Epigenetics. Chromatin. 8, 6.

Phipson,B., Lee,S., Majewski,I.J., Alexander,W.S., and Smyth,G.K. (2016). ROBUST HYPERPARAMETER ESTIMATION PROTECTS AGAINST HYPERVARIABLE GENES AND IMPROVES POWER TO DETECT DIFFERENTIAL EXPRESSION. Ann. Appl. Stat. 10, 946-963.

Poulikakos,P.I., Persaud,Y., Janakiraman,M., Kong,X., Ng,C., Moriceau,G., Shi,H., Atefi,M., Titz,B., Gabay,M.T., Salton,M., Dahlman,K.B., Tadi,M., Wargo,J.A., Flaherty,K.T., Kelley,M.C., Misteli,T., Chapman,P.B., Sosman,J.A., Graeber,T.G., Ribas,A., Lo,R.S., Rosen,N., and Solit,D.B. (2011). RAF inhibitor resistance is mediated by dimerization of aberrantly spliced BRAF(V600E). Nature 480, 387-390.

Rambow,F., Job,B., Petit,V., Gesbert,F., Delmas,V., Seberg,H., Meurice,G., Van,O.E., Dessen,P., Robert,C., Gautheret,D., Cornell,R.A., Sarasin,A., and Larue,L. (2015). New Functional Signatures for Understanding Melanoma Biology from Tumor Cell Lineage-Specific Analysis. Cell Rep. 13, 840-853.

Rambow,F., Rogiers,A., Marin-Bejar,O., Aibar,S., Femel,J., Dewaele,M., Karras,P., Brown,D., Chang,Y.H., biec-Rychter,M., Adriaens,C., Radaelli,E., Wolter,P., Bechter,O., Dummer,R., Levesque,M., Piris,A., Frederick,D.T., Boland,G., Flaherty,K.T., van den,O.J., Voet,T., Aerts,S., Lund,A.W., and Marine,J.C. (2018). Toward Minimal Residual Disease-Directed Therapy in Melanoma. Cell 174, 843-855.

Rastrelli,M., Tropea,S., Rossi,C.R., and Alaibac,M. (2014). Melanoma: epidemiology, risk factors, pathogenesis, diagnosis and classification. In Vivo 28, 1005-1011.

Ratterman,M., Hallmeyer,S., and Richards,J. (2016). Sequencing of New and Old Therapies for Metastatic Melanoma. Curr. Treat. Options. Oncol 17, 52.

Reczko,M., Maragkakis,M., Alexiou,P., Grosse,I., and Hatzigeorgiou,A.G. (2012). Functional microRNA targets in protein coding sequences. Bioinformatics 28, 771-776.

Reich,M., Liefeld,T., Gould,J., Lerner,J., Tamayo,P., and Mesirov,J.P. (2006). GenePattern 2.0. Nat. Genet. 38, 500-501.

Reimand,J., Arak,T., Adler,P., Kolberg,L., Reisberg,S., Peterson,H., and Vilo,J. (2016). g:Profiler-a web server for functional interpretation of gene lists (2016 update). Nucleic Acids Res 44, W83-W89.

Riesenberg,S., Groetchen,A., Siddaway,R., Bald,T., Reinhardt,J., Smorra,D., Kohlmeyer,J., Renn,M., Phung,B., Aymans,P., Schmidt,T., Hornung,V., Davidson,I., Goding,C.R., Jonsson,G., Landsberg,J., Tuting,T., and Holzel,M. (2015). MITF and c-Jun antagonism interconnects melanoma dedifferentiation with pro-inflammatory cytokine responsiveness and myeloid cell recruitment. Nat Commun. 6, 8755.

Rizos,H., Menzies,A.M., Pupo,G.M., Carlino,M.S., Fung,C., Hyman,J., Haydu,L.E., Mijatov,B., Becker,T.M., Boyd,S.C., Howle,J., Saw,R., Thompson,J.F., Kefford,R.F., Scolyer,R.A., and Long,G.V. (2014). BRAF inhibitor resistance mechanisms in metastatic melanoma: spectrum and clinical impact. Clin Cancer Res 20, 1965-1977.

Robinson,M.D., McCarthy,D.J., and Smyth,G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139-140.

Robinson,M.D. and Oshlack,A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol 11, R25.

Rosa,A.M., Dabas,N., Byrnes,D.M., Eller,M.S., and Grichnik,J.M. (2012). Germ cell proteins in melanoma: prognosis, diagnosis, treatment, and theories on expression. J Skin Cancer 2012, 621968.

Rosenthal,R., McGranahan,N., Herrero,J., Taylor,B.S., and Swanton,C. (2016). DeconstructSigs: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. Genome Biol 17, 31.

Ross,M.I. and Gershenwald,J.E. (2011). Evidence-based treatment of early-stage melanoma. J Surg Oncol 104, 341-353.

Sanborn,J.Z., Chung,J., Purdom,E., Wang,N.J., Kakavand,H., Wilmott,J.S., Butler,T., Thompson,J.F., Mann,G.J., Haydu,L.E., Saw,R.P., Busam,K.J., Lo,R.S., Collisson,E.A., Hur,J.S., Spellman,P.T., Cleaver,J.E., Gray,J.W., Huh,N., Murali,R., Scolyer,R.A., Bastian,B.C., and Cho,R.J. (2015). Phylogenetic analyses of melanoma reveal complex patterns of metastatic dissemination. Proc. Natl. Acad. Sci U. S. A 112, 10995-11000.

Sanders,D.S., Blessing,K., Hassan,G.A., Bruton,R., Marsden,J.R., and Jankowski,J. (1999). Alterations in cadherin and catenin expression during the biological progression of melanocytic tumours. Mol Pathol. 52, 151-157.

Schachter,J., Ribas,A., Long,G.V., Arance,A., Grob,J.J., Mortier,L., Daud,A., Carlino,M.S., McNeil,C., Lotem,M., Larkin,J., Lorigan,P., Neyns,B., Blank,C., Petrella,T.M., Hamid,O., Zhou,H., Ebbinghaus,S., Ibrahim,N., and Robert,C. (2017). Pembrolizumab versus ipilimumab for advanced melanoma: final overall survival results of a multicentre, randomised, open-label phase 3 study (KEYNOTE-006). Lancet 390, 1853-1862.

Schramm,S.J., Campain,A.E., Scolyer,R.A., Yang,Y.H., and Mann,G.J. (2012). Review and cross-validation of gene expression signatures and melanoma prognosis. J Invest Dermatol 132, 274-283.

Seashore-Ludlow,B., Rees,M.G., Cheah,J.H., Cokol,M., Price,E.V., Coletti,M.E., Jones,V., Bodycombe,N.E., Soule,C.K., Gould,J., Alexander,B., Li,A., Montgomery,P., Wawer,M.J., Kuru,N., Kotz,J.D., Hon,C.S., Munoz,B., Liefeld,T., Dancik,V., Bittker,J.A., Palmer,M., Bradner,J.E., Shamji,A.F., Clemons,P.A., and Schreiber,S.L. (2015). Harnessing Connectivity in a Large-Scale Small-Molecule Sensitivity Dataset. Cancer Discov. 5, 1210-1223.

Sensi,M., Catani,M., Castellano,G., Nicolini,G., Alciato,F., Tragni,G., De,S.G., Bersani,I., Avanzi,G., Tomassetti,A., Canevari,S., and Anichini,A. (2011). Human cutaneous melanomas lacking MITF and melanocyte differentiation antigens express a functional Axl receptor kinase. J Invest Dermatol 131, 2448-2457.

Sergushichev,A. (2016). An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. bioRxiv.

Serrone,L., Zeuli,M., Sega,F.M., and Cognetti,F. (2000). Dacarbazine-based chemotherapy for metastatic melanoma: thirty-year experience overview. J Exp. Clin Cancer Res 19, 21-34.

Shain,A.H. and Bastian,B.C. (2016). From melanocytes to melanomas. Nat Rev. Cancer 16, 345-358.

Shi,H., Moriceau,G., Kong,X., Lee,M.K., Lee,H., Koya,R.C., Ng,C., Chodon,T., Scolyer,R.A., Dahlman,K.B., Sosman,J.A., Kefford,R.F., Long,G.V., Nelson,S.F., Ribas,A., and Lo,R.S. (2012). Melanoma whole-exome sequencing identifies (V600E)B-RAF amplification-mediated acquired B-RAF inhibitor resistance. Nat Commun. 3, 724.

Sosman,J.A., Kim,K.B., Schuchter,L., Gonzalez,R., Pavlick,A.C., Weber,J.S., McArthur,G.A., Hutson,T.E., Moschos,S.J., Flaherty,K.T., Hersey,P., Kefford,R., Lawrence,D., Puzanov,I., Lewis,K.D., Amaravadi,R.K., Chmielowski,B., Lawrence,H.J., Shyr,Y., Ye,F., Li,J., Nolop,K.B., Lee,R.J., Joe,A.K., and Ribas,A. (2012). Survival in BRAF V600-mutant advanced melanoma treated with vemurafenib. N. Engl. J Med. 366, 707-714.

Sosman,J.A., Moon,J., Tuthill,R.J., Warneke,J.A., Vetto,J.T., Redman,B.G., Liu,P.Y., Unger,J.M., Flaherty,L.E., and Sondak,V.K. (2011). A phase 2 trial of complete resection for stage IV melanoma: results of Southwest Oncology Group Clinical Trial S9430. Cancer 117, 4740-06.

Steegmaier,M., Hoffmann,M., Baum,A., Lenart,P., Petronczki,M., Krssak,M., Gurtler,U., Garin-Chesa,P., Lieb,S., Quant,J., Grauert,M., Adolf,G.R., Kraut,N., Peters,J.M., and Rettig,W.J. (2007). BI 2536, a potent and selective inhibitor of polo-like kinase 1, inhibits tumor growth in vivo. Curr. Biol 17, 316-322.

Stern,R.S. (2001). The risk of melanoma in association with long-term exposure to PUVA. J Am Acad. Dermatol 44, 755-761.

Straussman,R., Morikawa,T., Shee,K., Barzily-Rokni,M., Qian,Z.R., Du,J., Davis,A., Mongare,M.M., Gould,J., Frederick,D.T., Cooper,Z.A., Chapman,P.B., Solit,D.B., Ribas,A., Lo,R.S., Flaherty,K.T., Ogino,S., Wargo,J.A., and Golub,T.R. (2012). Tumour micro-environment elicits innate resistance to RAF inhibitors through HGF secretion. Nature 487, 500-504.

Streicher,K.L., Zhu,W., Lehmann,K.P., Georgantas,R.W., Morehouse,C.A., Brohawn,P., Carrasco,R.A., Xiao,Z., Tice,D.A., Higgs,B.W., Richman,L., Jallal,B., Ranade,K., and Yao,Y. (2012). A novel oncogenic role for the miRNA-506-514 cluster in initiating melanocyte transformation and promoting melanoma growth. Oncogene 31, 1558-1570.

Sturm,R.A. (1998). Human pigmentation genes and their response to solar UV radiation. Mutat. Res 422, 69-76.

Sturm,R.A. (2002). Skin colour and skin cancer - MC1R, the genetic link. Melanoma Res 12, 405-416.

Subramanian,A., Tamayo,P., Mootha,V.K., Mukherjee,S., Ebert,B.L., Gillette,M.A., Paulovich,A., Pomeroy,S.L., Golub,T.R., Lander,E.S., and Mesirov,J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci U. S. A 102, 15545-15550.

Swope,V., Alexander,C., Starner,R., Schwemberger,S., Babcock,G., and bdel-Malek,Z.A. (2014). Significance of the melanocortin 1 receptor in the DNA damage response of human melanocytes to ultraviolet radiation. Pigment Cell Melanoma Res 27, 601-610.

The Cancer Genome Atlas Network (2015). Genomic Classification of Cutaneous Melanoma. Cell 161, 1681-1696.

Tirosh,I., Izar,B., Prakadan,S.M., Wadsworth,M.H., Treacy,D., Trombetta,J.J., Rotem,A., Rodman,C., Lian,C., Murphy,G., Fallahi-Sichani,M., Dutton-Regester,K., Lin,J.R., Cohen,O., Shah,P., Lu,D., Genshaft,A.S., Hughes,T.K., Ziegler,C.G., Kazer,S.W., Gaillard,A., Kolb,K.E., Villani,A.C., Johannessen,C.M., Andreev,A.Y., Van Allen,E.M., Bertagnolli,M., Sorger,P.K., Sullivan,R.J., Flaherty,K.T., Frederick,D.T., Jane-Valbuena,J., Yoon,C.H., Rozenblatt-Rosen,O., Shalek,A.K., Regev,A., and Garraway,L.A. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science 352, 189-196.

Titus-Ernstoff,L., Perry,A.E., Spencer,S.K., Gibson,J.J., Cole,B.F., and Ernstoff,M.S. (2005). Pigmentary characteristics and moles in relation to melanoma risk. Int J Cancer 116, 144-149.

Trunzer,K., Pavlick,A.C., Schuchter,L., Gonzalez,R., McArthur,G.A., Hutson,T.E., Moschos,S.J., Flaherty,K.T., Kim,K.B., Weber,J.S., Hersey,P., Long,G.V., Lawrence,D., Ott,P.A., Amaravadi,R.K., Lewis,K.D., Puzanov,I., Lo,R.S., Koehler,A., Kockx,M., Spleiss,O., Schell-Steven,A., Gilbert,H.N., Cockey,L., Bollag,G., Lee,R.J., Joe,A.K., Sosman,J.A., and Ribas,A. (2013). Pharmacodynamic effects and mechanisms of resistance to vemurafenib in patients with metastatic melanoma. J Clin Oncol 31, 1767-1774.

Tsao,H. and Niendorf,K. (2004). Genetic testing in hereditary melanoma. J Am Acad. Dermatol 51, 803-808.

Tsoi,J., Robert,L., Paraiso,K., Galvan,C., Sheu,K.M., Lay,J., Wong,D.J.L., Atefi,M., Shirazi,R., Wang,X., Braas,D., Grasso,C.S., Palaskas,N., Ribas,A., and Graeber,T.G. (2018). Multi-stage Differentiation Defines Melanoma Subtypes with Differential Vulnerability to Drug-Induced Iron-Dependent Oxidative Stress. Cancer Cell 33, 890-904.

Van Allen,E.M., Wagle,N., Sucker,A., Treacy,D.J., Johannessen,C.M., Goetz,E.M., Place,C.S., Taylor-Weiner,A., Whittaker,S., Kryukov,G.V., Hodis,E., Rosenberg,M., McKenna,A., Cibulskis,K., Farlow,D., Zimmer,L., Hillen,U., Gutzmer,R., Goldinger,S.M., Ugurel,S., Gogas,H.J., Egberts,F., Berking,C., Trefzer,U., Loquai,C., Weide,B., Hassel,J.C., Gabriel,S.B., Carter,S.L., Getz,G., Garraway,L.A., and Schadendorf,D. (2014). The genetic landscape of clinical resistance to RAF inhibition in metastatic melanoma. Cancer Discov. 4, 94-109.

van Jaarsveld,R.H. and Kops,G.J.P.L. (2016). Difference Makers: Chromosomal Instability versus Aneuploidy in Cancer. Trends Cancer 2, 561-571.

Veierod,M.B., Adami,H.O., Lund,E., Armstrong,B.K., and Weiderpass,E. (2010). Sun and solarium exposure and melanoma risk: effects of age, pigmentary characteristics, and nevi. Cancer Epidemiol. Biomarkers Prev. 19, 111-120.

Verfaillie,A., Imrichova,H., Atak,Z.K., Dewaele,M., Rambow,F., Hulselmans,G., Christiaens,V., Svetlichnyy,D., Luciani,F., Van den,M.L., Claerhout,S., Fiers,M., Journe,F., Ghanem,G.E., Herrmann,C., Halder,G., Marine,J.C., and Aerts,S. (2015). Decoding the regulatory landscape of melanoma reveals TEADS as regulators of the invasive cell state. Nat. Commun. 6, 6683.

Villanueva,J., Vultur,A., Lee,J.T., Somasundaram,R., Fukunaga-Kalabis,M., Cipolla,A.K., Wubbenhorst,B., Xu,X., Gimotty,P.A., Kee,D., Santiago-Walker,A.E., Letrero,R., D'Andrea,K., Pushparajan,A., Hayden,J.E., Brown,K.D., Laquerre,S., McArthur,G.A., Sosman,J.A., Nathanson,K.L., and Herlyn,M. (2010). Acquired resistance to BRAF inhibitors mediated by a RAF kinase switch in melanoma can be overcome by cotargeting MEK and IGF-1R/PI3K. Cancer Cell 18, 683-695.

Wagle,N., Emery,C., Berger,M.F., Davis,M.J., Sawyer,A., Pochanard,P., Kehoe,S.M., Johannessen,C.M., Macconaill,L.E., Hahn,W.C., Meyerson,M., and Garraway,L.A. (2011). Dissecting therapeutic resistance to RAF inhibition in melanoma by tumor genomic profiling. J Clin Oncol 29, 3085-3096.

Wang,X. (2008). miRDB: a microRNA target prediction and functional annotation database with a wiki interface. RNA. 14, 1012-1017.

Wei,I.H., Healy,M.A., and Wong,S.L. (2014). Surgical treatment options for stage IV melanoma. Surg Clin North Am 94, 1075-89, ix.

Whiteman,D.C., Whiteman,C.A., and Green,A.C. (2001). Childhood sun exposure as a risk factor for melanoma: a systematic review of epidemiologic studies. Cancer Causes Control 12, 69-82.

Whittaker,S.R., Theurillat,J.P., Van,A.E., Wagle,N., Hsiao,J., Cowley,G.S., Schadendorf,D., Root,D.E., and Garraway,L.A. (2013). A genome-scale RNA interference screen implicates NF1 loss in resistance to RAF inhibition. Cancer Discov. 3, 350-362.

Widmer,D.S., Cheng,P.F., Eichhoff,O.M., Belloni,B.C., Zipser,M.C., Schlegel,N.C., Javelaud,D., Mauviel,A., Dummer,R., and Hoek,K.S. (2012). Systematic classification of melanoma cells by phenotype-specific gene expression mapping. Pigment Cell Melanoma Res 25, 343-353.

Wilkerson,M.D. and Hayes,D.N. (2010). ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics 26, 1572-1573.

Winnepenninckx,V., Lazar,V., Michiels,S., Dessen,P., Stas,M., Alonso,S.R., Avril,M.F., Ortiz Romero,P.L., Robert,T., Balacescu,O., Eggermont,A.M., Lenoir,G., Sarasin,A., Tursz,T., van den Oord,J.J., and Spatz,A. (2006). Gene expression profiling of primary cutaneous melanoma and clinical outcome. J Natl. Cancer Inst. 98, 472-482.

Wolchok,J.D., Chiarion-Sileni,V., Gonzalez,R., Rutkowski,P., Grob,J.J., Cowey,C.L., Lao,C.D., Wagstaff,J., Schadendorf,D., Ferrucci,P.F., Smylie,M., Dummer,R., Hill,A., Hogg,D., Haanen,J., Carlino,M.S., Bechter,O., Maio,M., Marquez-Rodas,I., Guidoboni,M., McArthur,G., Lebbe,C., Ascierto,P.A., Long,G.V., Cebon,J., Sosman,J., Postow,M.A., Callahan,M.K., Walker,D., Rollin,L., Bhore,R., Hodi,F.S., and Larkin,J. (2017). Overall Survival with Combined Nivolumab and Ipilimumab in Advanced Melanoma. N. Engl. J Med. 377, 1345-1356.

Wong,S.L., Faries,M.B., Kennedy,E.B., Agarwala,S.S., Akhurst,T.J., Ariyan,C., Balch,C.M., Berman,B.S., Cochran,A., Delman,K.A., Gorman,M., Kirkwood,J.M., Moncrieff,M.D., Zager,J.S., and Lyman,G.H. (2018). Sentinel Lymph Node Biopsy and Management of

Regional Lymph Nodes in Melanoma: American Society of Clinical Oncology and Society of Surgical Oncology Clinical Practice Guideline Update. J Clin Oncol 36, 399-413.

Wouters,J., Stas,M., Govaere,O., Barrette,K., Dudek,A., Vankelecom,H., Haydu,L.E., Thompson,J.F., Scolyer,R.A., and van den Oord,J.J. (2014). A novel hypoxia-associated subset of FN1 high MITF low melanoma cells: identification, characterization, and prognostic value. Mod. Pathol. 27, 1088-1100.

Yancovitz,M., Litterman,A., Yoon,J., Ng,E., Shapiro,R.L., Berman,R.S., Pavlick,A.C., Darvishian,F., Christos,P., Mazumdar,M., Osman,I., and Polsky,D. (2012). Intra- and inter-tumor heterogeneity of BRAF(V600E) mutations in primary and metastatic melanoma. PLoS One 7, e29336.

Yokoyama,S., Feige,E., Poling,L.L., Levy,C., Widlund,H.R., Khaled,M., Kung,A.L., and Fisher,D.E. (2008). Pharmacologic suppression of MITF expression via HDAC inhibitors in the melanocyte lineage. Pigment Cell Melanoma Res 21, 457-463.

Yoshihara,K., Shahmoradgoli,M., Martinez,E., Vegesna,R., Kim,H., Torres-Garcia,W., Trevino,V., Shen,H., Laird,P.W., Levine,D.A., Carter,S.L., Getz,G., Stemke-Hale,K., Mills,G.B., and Verhaak,R.G. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. Nat. Commun. 4, 2612.

Yoshimura,K., Meckel,K.F., Laird,L.S., Chia,C.Y., Park,J.J., Olino,K.L., Tsunedomi,R., Harada,T., Iizuka,N., Hazama,S., Kato,Y., Keller,J.W., Thompson,J.M., Chang,F., Romer,L.H., Jain,A., Iacobuzio-Donahue,C., Oka,M., Pardoll,D.M., and Schulick,R.D. (2009). Integrin alpha2 mediates selective metastasis to the liver. Cancer Res 69, 7320-7328.

Yu,L., Lu,M., Jia,D., Ma,J., Ben-Jacob,E., Levine,H., Kaipparettu,B.A., and Onuchic,J.N. (2017). Modeling the Genetic Regulation of Cancer Metabolism: Interplay between Glycolysis and Oxidative Phosphorylation. Cancer Res 77, 1564-1574.

Zhang,M., Qureshi,A.A., Geller,A.C., Frazier,L., Hunter,D.J., and Han,J. (2012). Use of tanning beds and incidence of skin cancer. J Clin Oncol 30, 1588-1593.

Zipser,M.C., Eichhoff,O.M., Widmer,D.S., Schlegel,N.C., Schoenewolf,N.L., Stuart,D., Liu,W., Gardner,H., Smith,P.D., Nuciforo,P., Dummer,R., and Hoek,K.S. (2011). A proliferative melanoma cell phenotype is responsive to RAF/MEK inhibition independent of BRAF mutation status. Pigment Cell Melanoma Res 24, 326-333.

8 LIST OF PUBLICATIONS

Below is the list of publications to which I contributed with my bioinformatic skills. Publications in the melanoma filed are underlined.

1. Huber V, Vallacchi V, Fleming V, Hu X, Cova A, **Dugo M**, Shahaj E, Sulsenti R, Vergani E, Filipazzi P, De Laurentiis A, Lalli L, Di Guardo L, Patuzzo R, Vergani B, Casiraghi E, Cossa M, Gualeni A, Bollati V, Arienti F, De Braud F, Mariani L, Villa A, Altevogt P, Umansky V, Rodolfo M, Rivoltini L. *Tumor-derived microRNAs induce myeloid suppressor cells and predict immunotherapy resistance in melanoma*. J Clin Invest. 2018 Sep 27.

2. Castagnoli L, Iorio E, **Dugo M**, Koschorke A, Faraci S, Canese R, Casalini P, Nanni P, Vernieri C, Di Nicola M, Morelli D, Tagliabue E, Pupa SM. *Intratumor lactate levels reflect HER2 addiction status in HER2-positive breast cancer*. J Cell Physiol. 2018 Aug 21.

3. **Dugo M**, Huang X, Iorio MV, Cataldo A, Tagliabue E, Daidone MG, Wu J, Orlandi R. *MicroRNA co-expression patterns unravel the relevance of extra cellular matrix and immunity in breast cancer*. Breast. 2018 Jun;39:46-52.

4. Andriani F, Majorini MT, Mano M, Landoni E, Miceli R, Facchinetti F, Mensah M, Fontanella E, **Dugo M**, Giacca M, Pastorino U, Sozzi G, Delia D, Roz L, Lecis D. *MiR-16 regulates the pro-tumorigenic potential of lung fibroblasts through the inhibition of HGF production in an FGFR-1- and MEK1-dependent manner*. J Hematol Oncol. 2018 Mar 20;11(1):45.

5. Romeo P, Colombo C, Granata R, Calareso G, Gualeni AV, **Dugo M**, De Cecco L, Rizzetti MG, Zanframundo A, Aiello A, Carcangiu ML, Gloghini A, Ferrero S, Licitra L, Greco A, Fugazzola L, Locati LD, Borrello MG. *Circulating miR-375 as a novel prognostic marker for metastatic medullary thyroid cancer patients*. Endocr Relat Cancer. 2018 Mar;25(3):217-231.

6. Callari M, **Dugo M**, Miodini P, Veneroni S, Bianchini G, Daidone MG, Cappelletti V. *Dissecting Time- from Tumor-Related Gene Expression Variability in Bilateral Breast Cancer*. Int J Mol Sci. 2018 Jan 9;19(1). pii: E196.

7. Gregorini M, Corradetti V, Pattonieri EF, Rocca C, Milanesi S, Peloso A, Canevari S, De Cecco L, **Dugo M**, Avanzini MA, Mantelli M, Maestri M, Esposito P, Bruno S, Libetta C, Dal Canton A, Rampino T. *Perfusion of isolated rat kidney with Mesenchymal Stromal Cells/Extracellular Vesicles prevents ischaemic injury*. J Cell Mol Med. 2017 Dec;21(12):3381-3393.

8. El Bezawy R, Cominetti D, Fenderico N, Zuco V, Beretta GL, **Dugo M**, Arrighetti N, Stucchi C, Rancati T, Valdagni R, Zaffaroni N, Gandellini P. *miR-875-5p counteracts epithelial-to-mesenchymal transition and enhances radiation response in prostate cancer through repression of the EGFR-ZEB1 axis*. Cancer Lett. 2017 Jun 1;395:53-62.

9. Verri C, Borzi C, Holscher T, **Dugo M**, Devecchi A, Drake K, Sestini S, Suatoni P, Romeo E, Sozzi G, Pastorino U, Boeri M. *Mutational Profile from Targeted NGS Predicts Survival in LDCT Screening-Detected Lung Cancers*. J Thorac Oncol. 2017 Jun;12(6):922-931.

10. Castagnoli L, Ghedini GC, Koschorke A, Triulzi T, **Dugo M**, Gasparini P, Casalini P, Palladini A, Iezzi M, Lamolinara A, Lollini PL, Nanni P, Chiodoni C, Tagliabue E, Pupa SM. *Pathobiological implications of the d16HER2 splice variant for stemness and aggressiveness of HER2-positive breast cancer*. Oncogene. 2017 Mar 23;36(12):1721-1732.

11. Pintarelli G, Cotroneo CE, Noci S, **Dugo M**, Galvan A, Delli Carpini S, Citterio L, Manunta P, Incarbone M, Tosi D, Santambrogio L, Dragani TA, Colombo F. *Genetic susceptibility variants for lung cancer: replication study and assessment as expression quantitative trait loci*. Sci Rep. 2017 Feb 9;7:42185.

12. Cappelletti V, Iorio E, Miodini P, Silvestri M, **Dugo M**, Daidone MG. *Metabolic Footprints and Molecular Subtypes in Breast Cancer*. Dis Markers. 2017;2017:7687851.

13. Vallacchi V, Camisaschi C, **Dugo M**, Vergani E, Deho P, Gualeni A, Huber V, Gloghini A, Maurichi A, Santinami M, Sensi M, Castelli C, Rivoltini L, Rodolfo M. *microRNA Expression in Sentinel Nodes from Progressing Melanoma Patients Identifies Networks Associated with Dysfunctional Immune Response*. Genes (Basel). 2016 Dec 14;7(12).

14. Veneroni S, **Dugo M**, Daidone MG, Iorio E, Valeri B, Pinciroli P, De Bortoli M, Marchesi E, Miodini P, Taverna E, Ricci A, Canevari S, Pelosi G, Bongarzone I. *Applicability of Under Vacuum Fresh Tissue Sealing and Cooling to Omics Analysis of Tumor Tissues*. Biopreserv Biobank. 2016 Dec;14(6):480-490.

15. **Dugo M**, Cotroneo CE, Lavoie-Charland E, Incarbone M, Santambrogio L, Rosso L, van den Berge M, Nickle D, Paré PD, Bossé Y, Dragani TA, Colombo F. *Human Lung Tissue Transcriptome: Influence of Sex and Age*. PLoS One. 2016 Nov 30;11(11):e0167460.

16. Gimondi S, **Dugo M**, Vendramin A, Bermema A, Biancon G, Cavané A, Corradini P, Carniti C. Circulating miRNA panel for prediction of acute graft-versus-host disease in lymphoma patients undergoing matched unrelated hematopoietic stem cell transplantation. Exp Hematol. 2016 Jul;44(7):624-634.e1.

17. Cotroneo CE, Dassano A, Colombo F, Pettinicchio A, Lecis D, **Dugo M**, De Cecco L, Dragani TA, Manenti G. *Expression quantitative trait analysis reveals fine germline transcript regulation in mouse lung tumors*. Cancer Lett. 2016 Jun 1;375(2):221-230.

18. Merlino G, Miodini P, Paolini B, Carcangiu ML, Gennaro M, **Dugo M**, Daidone MG, Cappelletti V. *Stromal Activation by Tumor Cells: An in Vitro Study in Breast Cancer*. Microarrays (Basel). 2016 May 18;5(2).

19. Minna E, Romeo P, **Dugo M**, De Cecco L, Todoerti K, Pilotti S, Perrone F, Seregni E, Agnelli L, Neri A, Greco A, Borrello MG. *miR-451a is underexpressed and targets AKT/mTOR pathway in papillary thyroid carcinoma*. Oncotarget. 2016 Mar 15;7(11):12731-47.

20. Mauri G, Jachetti E, Comuzzi B, **Dugo M**, Arioli I, Miotti S, Sangaletti S, Di Carlo E, Tripodo C, Colombo MP. Genetic deletion of osteopontin in TRAMP mice skews prostate carcinogenesis from adenocarcinoma to aggressive human-like neuroendocrine cancers. Oncotarget. 2016 Jan 26;7(4):3905-20.

21. Vergani E, Di Guardo L, **Dugo M**, Rigoletto S, Tragni G, Ruggeri R, Perrone F, Tamborini E, Gloghini A, Arienti F, Vergani B, Deho P, De Cecco L, Vallacchi V, Frati P, Shahaj E, Villa A, Santinami M, De Braud F, Rivoltini L, Rodolfo M. *Overcoming melanoma resistance to vemurafenib by targeting CCL2-induced miR-34a, miR-100 and miR-125b*. Oncotarget. 2016 Jan 26;7(4):4428-41.

22. Martinez-Lozano Sinues P, Landoni E, Miceli R, Dibari VF, **Dugo M**, Agresti R, Tagliabue E, Cristoni S, Orlandi R. Secondary electrospray ionization-mass spectrometry and a novel statistical bioinformatic approach identifies a cancer-related profile in exhaled breath of breast cancer patients: a pilot study. J Breath Res. 2015 Sep 21;9(3):031001.

23. Huang X, **Dugo M**, Callari M, Sandri M, De Cecco L, Valeri B, Carcangiu ML, Xue J, Bi R, Veneroni S, Daidone MG, Ménard S, Tagliabue E, Shao Z, Wu J, Orlandi R. *Molecular portrait of breast cancer in China reveals comprehensive transcriptomic likeness to Caucasian breast cancer and low prevalence of luminal A subtype*. Cancer Med. 2015 Jul;4(7):1016-30.

24. Musella V, Callari M, Di Buduo E, Scuro M, **Dugo M**, Miodini P, Bianchini G, Paolini B, Gianni L, Daidone MG, Cappelletti V. *Use of formalin-fixed paraffin-embedded samples for gene expression studies in breast cancer patients*. PLoS One. 2015 Apr 6;10(4):e0123194.

25. **Dugo M**, Nicolini G, Tragni G, Bersani I, Tomassetti A, Colonna V, Del Vecchio M, De Braud F, Canevari S, Anichini A, Sensi M. A melanoma subtype with intrinsic resistance to BRAF inhibition identified by receptor tyrosine kinases gene-driven classification. Oncotarget. 2015 Mar 10;6(7):5118-33.

26. Galvan A, Colombo F, Frullanti E, Dassano A, Noci S, Wang Y, Eisen T, Matakidou A, Tomasello L, Vezzalini M, Sorio C, **Dugo M**, Ambrogi F, Iacobucci I, Martinelli G, Incarbone M, Alloisio M, Nosotti M, Tosi D, Santambrogio L, Pelosi G, Pastorino U, Houlston RS, Dragani TA. *Germline polymorphisms and survival of lung adenocarcinoma patients: a genome-wide study in two European patient series*. Int J Cancer. 2015 Mar 1;136(5):E262-71.